

DOCUMENT RESUME

ED 283 028

CE 047 500

AUTHOR Branch, Alvia Y.; And Others
TITLE Summer Training and Education Program (STEP). Report on the 1985 Summer Experience.
INSTITUTION Public/Private Ventures, Philadelphia, PA.
PUB DATE Apr 86
NOTE 174p.; For the 1986 report, see CE 047 499.
AVAILABLE FROM Public/Private Ventures, 399 Market Street, Philadelphia, PA 19106.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Attitude Change; *Basic Skills; Comparative Analysis; Demonstration Programs; *Disadvantaged Youth; *Economically Disadvantaged; Geographic Location; National Programs; Outcomes of Education; Pretests Posttests; Program Effectiveness; *Remedial Instruction; Secondary Education; *Summer Programs; *Youth Programs

IDENTIFIERS Impact Studies; *Summer Training and Education Program

ABSTRACT

The summer of 1985 marked the first full-scale implementation of the Summer Training and Education Program (STEP) model. Federally funded, STEP provides 14- and 15-year-olds with a paid summer experience that combines work, remediation in reading and math, and life skills instruction with an emphasis on responsible social and sexual attitudes and behavior. Participants spend two consecutive summers in the program. The first cohort consisted of 1,592 youths (each receiving \$3.35 per hour). They were randomly assigned to treatment and control groups at five demonstration sites--Boston, Fresno, San Diego, Seattle, and Portland, Oregon. In the posttest, the youths in the treatment group outscored their control group counterparts in both reading and math by approximately one quarter of a grade equivalent. The impact of participation varied as a function of sex, race, and site. Girls in the treatment group showed greater reading gains over their counterparts in the control group than did the boys. Sex was not as great a factor in the math scores, however. Hispanic youths showed the greatest gains, with Asians and Blacks deriving solid benefits as well. The impact of participation varied among the sites, with the most impact occurring in Boston and the least occurring in Portland. The program also appeared to result in changes in self-reported sexual behavior of boys, an effect unusual among pregnancy prevention programs. (Appendixes include technical data tables and discussions of the test floor effect, attrition, and control group losses.) (MN)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED283028

Public/Private Ventures

399 Market Street
Philadelphia, PA 19106
(215) 592-9099

Summer Training and Education Program (STEP)

Report on the 1985 Summer Experience

by Alvia Y. Branch
Julita Milliner
Jon Bumbaugh

April 1986

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
- ☐ Minor changes have been made to improve
reproduction quality.

- Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL IN MICROFICHE ONLY
HAS BEEN GRANTED BY

N. Jaffe

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Public/Private Ventures Board of Directors

Rex D. Adams
Vice President
Employee Relations
Mobil Oil Corporation
New York, New York

Judy Barker
President
The Borden Foundation
Columbus, Ohio

Alan K. Campbell
Vice Chairman
ARA Services, Inc.
Philadelphia, Pennsylvania

Alonzo A. Crim
Superintendent of Schools
Atlanta, Georgia

Alice F. Emerson
President
Wheaton College
Norton, Massachusetts

Sol Hurwitz
Senior Vice President
Committee for Economic
Development
New York, New York

Katharine C. Lyall
Executive Vice President
University of Wisconsin System
Madison, Wisconsin

Arabella Martinez
Arabella Martinez Institute
Berkeley, California

James L. Medoff
Professor, Department of
Economics
Harvard University
Cambridge, Massachusetts

Steven A. Minter
Director
The Cleveland Foundation
Cleveland, Ohio

Ersa H. Poston
Personnel Management
Consultant
McLean, Virginia

William J. Spring
Vice President
Federal Reserve Bank of Boston
Boston, Massachusetts

Mitchell Sviridoff
Professor, Graduate School of
Urban Management
New School for Social Research
New York, New York

Alfred M. Zuck
Executive Director
National Association of Schools
of Public Affairs and
Administration
Washington, D.C.

Alan K. Campbell
Chairman of the Board

Michael A. Bailin
President

Gary Walker
Executive Vice President

Public/Private Ventures is a national, not-for-profit corporation that designs, manages and evaluates social policy initiatives designed to help those whose lack of preparation for the workforce hampers their chances for productive lives. Its work is supported by funds from both the public and private sectors.

Further information about P/PV and its publications is available from: The Communications Department, Public/Private Ventures, 399 Market Street, Philadelphia, Pennsylvania 19106. Telephone: 215-592-9099.

ACKNOWLEDGEMENTS

The authors wish to acknowledge some of the many persons who made important contributions to this report. Wendy Wolf, Director of Research at P/PV, provided overall guidance and gave specific advice regarding analytical and research design issues. Dr. Andrew Sum provided an early review and comments on issues related to research design and analytic strategies. Richard Voith, Rickey Etheredge, Ken Libert, Eleanor Hammond and Batia Trietsch -- members of P/PV's data unit -- provided invaluable assistance with the management and analysis of the data. Michael Bailin, Michael Sack, Thomas J. Smith, Frances Vilella-Velez and Gary Walker reviewed early drafts of the report and contributed valuable insights. Paula Strawberry typed numerous drafts of the manuscript. Natalie Jaffe, Judy Wilmoth White and Ari Eisinger made important editorial contributions. Finally, many individuals at the STEP demonstration sites contributed both time and insights.

CONTENTS

	<u>Page</u>
EXECUTIVE SUMMARY	1
I. INTRODUCTION	1
Background	1
The STEP Model	3
Description of the Sites	8
Organization of This Report	16
II. RESEARCH DESIGN	17
Implementation Analysis	17
Impact Analysis	18
Overview of Design Strategies	19
Data Collection Strategy	21
III. THE IMPLEMENTATION OF STEP	25
The Participant Selection Process	25
The Work Experience Component	29
The Life Skills and Opportunities (LSO) Component	32
The Remediation Component	33
Summary	36
IV. STEP PARTICIPANT CHARACTERISTICS	39
Eligibility Requirements and Characteristics of the Sample	39
Other Demographic Characteristics	46
Additional Educational Characteristics	48
Fertility-Related Characteristics	50
Summary	54
V. PROGRAM IMPACT	55
Basic Analytic Strategies	55
Modeling Program Participation	59
Academic Performance	62
Return to School	81
Life Skills Outcomes	82
Fertility-Related Outcomes	83
VI. SUMMARY AND CONCLUSIONS	91
Program Implementation	91
Program Impact	93

CONTENTS (cont'd)

Page

APPENDICES

A.	Test Floor Effect	97
B.	Attrition	103
C.	Control Group Losses	109
D.	Technical Data Tables	115

LIST OF TABLES

	<u>Page</u>
1. Organizations Operating STEP, by Site	10
2. STEP Data Base, by Source	22
3. Number (Percent) of Youth Pre- and Post-tested	30
4. Distribution of STEP Youth, by Program Eligibility Criteria	41
5a. Distribution of STEP Youth's MAT Scores in Reading, by Selected Variables	44
5b. Distribution of STEP Youth's MAT Scores in Math, by Selected Variables	45
6. Demographic Characteristics of the STEP Sample at Baseline	47
7. Additional Educational Characteristics of the STEP Sample at Baseline	49
8. Fertility-Related Characteristics of STEP Youth at Baseline	52
9. Selected Characteristics of the Randomization, Learning Analysis and Life Skills Analysis Samples	58
10. Total Program Hours, by Site	60
11. Remediation Hours, by Site for Treatments Only	61
12. Determinants of Remediation Hours	63
13. Net Impact of Participation in STEP on MAT Scaled Scores in Reading and Math	64
14. Pre-test to Post-test Changes in Reading and Math Scores, by Group	66
15. Pre-test to Post-test Changes in Reading and Math Scores in San Diego, by Remediation Site	69
16. Net Impact of Participation in STEP as Revealed by Multivariate Analysis of MAT Scaled Scores in Reading and Math	72
17. Pre-test to Post-test Changes in Scores, by Race	74

LIST OF TABLES (cont'd)

	<u>Page</u>
18. Statistically Significant Impacts in Pooled Regression Model with Multiple Interaction Terms	76
19. Program Impact According to Pre-Program Achievement Levels	78
20. Math Impact with Treatment Interacted with Ability Levels in a Pooled Regression	80
21. Net Impact of Participation in STEP on Birth Control Knowledge Variables, by Site and Selected Subgroups ...	85
22. Percent of Respondents Sexually Active in the Previous Two Months	86

EXECUTIVE SUMMARY

Each year, large numbers of youth enter the labor market without basic skills. Their deficiencies are, in many cases, attributable to the lack of a high school diploma -- in itself a factor that reduces long-term employment and earnings prospects.

Research has identified poor academic performance and teenage pregnancy as important contributors to dropout behavior and subsequent difficulties in the labor market. Research has also noted that the early adolescent years are critical to a youth's school progress, and that the summers are times when disadvantaged youth tend to slow, halt and even regress in that progress. The programmatic vehicle selected for addressing these needs was the federally-funded Summer Youth Employment and Training Program (SYETP). SYETP is the largest national program that serves disadvantaged youth, and is operated in every locality in the United States. Incorporating remedial education and other instruction (as well as a school year follow-up) into SYETP produced the Summer Training and Education Program, a national demonstration designed and managed by Public/Private Ventures.

The STEP intervention provides 14- and 15-year-olds with a paid summer experience that combines work, remediation in reading and math, and life skills instruction that emphasizes responsible social and sexual attitudes and behavior. A summer 1984 pilot of the STEP model proved successful, and established the basic parameters of the model that was implemented in the full-scale demonstration initiated in the summer of 1985.

The first summer of the full demonstration yielded encouraging results:

- The STEP design proved operationally feasible and attracted good participation from the target group. A racially diverse group with multiple problems, these youth were clearly in need of an intervention like STEP by virtue of their low economic status, poor academic history and early involvement in sexual activity, which increases the risk of adolescent parenthood. Without an intervention like STEP, many of these youth are likely to drop out of school, experience limited success in the marketplace, and find themselves unable to support the families they will form.
- STEP was successful in stemming the learning losses that would have occurred in the absence of the treatment. While treatment youth scored significantly higher than control group youth at post-test in both reading and math, in the aggregate, both groups experienced losses. The control group losses were

significantly larger than those experienced by treatment youth. Some gains among subsets of treatment youth were also realized -- Asians and youth enrolled at three of San Diego's remediation sites showed pre-test to post-test gains.

- Although most treatment subgroups improved vis-a-vis their control group counterparts in either reading or math, girls and Hispanic youth appear to have benefited most from participation in STEP. Treatment girls outscored control group girls in reading by half a grade equivalent, and in math by one-quarter of a grade equivalent. Hispanic youth in the treatment group outscored Hispanic control group youth by half a grade equivalent or more in both reading and math.
- Participants' knowledge of birth control was substantially increased, and one behavioral impact was achieved: significantly more treatment boys than control group boys reported abstaining from sex during the program period.

The STEP demonstration is being managed by Public/Private Ventures (P/PV), an organization that designs, manages and evaluates social policy initiatives designed to improve job training and education for disadvantaged youth and adults. The demonstration is funded by public and private funds from both national and local sources, including local Job Training Partnership Act Title II-B funds, the U.S. Department of Labor, The Ford Foundation, The Edna McConnell Clark Foundation, The William and Flora Hewlett Foundation, James C. Penney Foundation, Aetna Life and Casualty Foundation, The Ahmanson Foundation and Lilly Endowment, Inc.

THE STEP MODEL

STEP operates within an existing, federally-funded program that serves large numbers of disadvantaged youth throughout the country during the summer months -- the Summer Youth Employment and Training Program (SYETP). It provides summer work experience and income for youth who otherwise would be unlikely to find jobs. Although SYETP has served many youth over the years, it has had only limited success in helping them make successful transitions to the world of work. STEP demonstrates some additional possibilities for this national program.

STEP's central objective is to provide participants with the means to stay in high school and earn a diploma. In the short term, STEP focuses on attaining objectives that can be observed within the course of a single summer. Among them are:

- increased achievement in reading and math or a reduction in the learning losses usually experienced in the summer months; and
- increased ability to make responsible decisions about sexual behavior.

The demonstration is being conducted in five sites: Boston, Massachusetts; Seattle, Washington; Portland, Oregon; Fresno, California; and San Diego, California. The sites were selected from a number of applicant sites based on their ability to operate an innovative program and to meet program and research design specifications.

Participants in STEP spend two consecutive summers in a program that provides both academic remediation and instruction in life skills in addition to the work experience traditionally available to them through SYETP. In the intervening school year, they take part in a voluntary support component designed to reinforce gains achieved during the first summer and to create expectations for additional gains in the future.

The first cohort of 1,593 14- and 15-year-olds -- half in a treatment group, and half in a control group -- was involved in the summer 1985 program. A second group of equal size will begin the program in 1986.

Four separate but related components comprise the STEP treatment: remediation, work experience, life skills and opportunities, and school year support.

The remediation component provides 90 hours of basic skills instruction in reading and math, with the aim of producing learning gains or, at least, decreasing the losses disadvantaged youth usually experience over the summer months. The STEP model calls for a remediation curriculum that is individualized, self-paced, and competency-based and that utilizes computer-assisted instruction (CAI). Curriculum guidelines mandate clearly-stated competencies, a comprehensive diagnostic test for reading and math, and instructional materials and assessment exercises that can be explicitly related to identified competencies.

The life skills and opportunities component presents a curriculum aimed at developing responsible social and sexual attitudes and behavior among participating youth. It covers such topics as personal decision-making, job equality issues, the consequences of teenage parenting and ways to avoid pregnancy, among them abstinence. The curriculum uses varied presentation formats including lectures, discussions, field trips, role-playing, films and outside speakers.

Work experience is provided through the federally-sponsored Summer Youth Employment and Training Program (SYETP), which funds entry-level jobs in clerical, maintenance and parks and recrea-

tion work. Youth are paid the minimum wage for their participation in both the work experience and remediation/life skills activities.

The school year support component seeks to ensure that treatment youth maintain the academic progress achieved during the summer and return to STEP for the second summer.

To qualify for participation in STEP, youth must be 14 or 15; satisfy the income requirements for participation in the Summer Youth Employment and Training Program, which stipulate membership in economically disadvantaged families; and be in need of academic remediation as demonstrated by low scores on standardized tests, a history of grade retention, or both.

The research design is an integral part of the STEP demonstration. It includes both short-term and long-term objectives, among them: assessment of the model's feasibility; measurement of participants' short-term educational gains and their development of responsible social and sexual attitudes and behavior; and an evaluation of the impact of STEP on high school retention and early labor force experiences. The research report, which presents an assessment of the first summer of the demonstration, focuses on STEP's short-term objectives.

A total of 1,593 STEP-eligible youth were randomly assigned to treatment and control group conditions in the five demonstration sites. Treatment youth participated in the STEP program; control group youth participated in regular work experience. Both were compensated at the same rate -- \$3.35 an hour.

In order to minimize the possibility that control group youth would have less motivation than treatment youth to perform well at post-test -- and thereby spuriously elevate the estimate of program impact -- an incentive scheme that followed a lottery format was devised to provide both groups with an external reason for doing their best. Any control group losses that emerge in spite of these incentives would presumably reflect true learning losses.

DEMONSTRATION FINDINGS

In the first two months of its 15-month treatment, STEP achieved many of its objectives. The demonstration was implemented in accordance with the basic program design and STEP's target population -- a racially diverse group of youth at risk of dropping out of high school -- was served as planned. Similarly, STEP was successful in attaining a number of benefits for its participants. Participants fared significantly better than their control group counterparts in reading and math. In addition, participation in STEP substantially increased participants' knowledge of birth control information, and more treatment boys than

control group boys reported abstaining from sex during the program period.

Program Implementation

STEP implementation required close collaboration among a variety of institutions at a level of intensity that was new to most sites. Employment and training agencies, school systems and community organizations collaborated in the implementation of STEP to an unusual degree. Local administrative arrangements varied widely, reflecting both the flexibility of the STEP model and traditional roles and responsibilities within sites. In all sites the level of commitment to STEP was high.

The demonstration recruited a large number of youth clearly in need of a complex, multiple treatment program such as that offered by STEP. By any criterion, this group of young people is highly at risk of dropping out of high school. At 14 and 15, many have already suffered significant academic setbacks. And, although most are sexually experienced, they are not very knowledgeable about contraception, nor do they practice it with any frequency.

Work experience, the most familiar component in the program mix, posed no major obstacles despite the challenge of finding half-time assignments and work appropriate for 14-year-olds.

Life skills instruction, although new to the sites, was delivered with a high degree of success. This seems due, in part, to a uniform curriculum of engaging content and format that provided teachers with a lesson plan, materials and informational resources.

Remediation, the newest and least tried component of the STEP model, was inevitably the most difficult for the sites to implement. At all sites, participants were offered 90 hours of concentrated reading and math instruction that employed a competency-based, self-paced and individualized curriculum and used computers to augment the instruction. In addition, attendance at the remediation sessions was quite high -- on average, 82 percent of all available hours.

For the remediation component, P/PV designed an approach called Community Organized Remedial Education (CORE) that delineated specific elements of the remediation program, detailed a development process for local educators and curriculum specialists to follow, and provided technical assistance to sites throughout the late spring. The pilot experience in 1984 had pointed conclusively to the advisability of such an approach, rather than one that imposed a pre-packaged curriculum on local educators or one that gave them such wide latitude that essential treatment uniformity across sites could not be ensured.

Nevertheless, significant problems were encountered in the implementation of this component, among them a late start-up, the lack of clear guidelines from P/PV regarding the content of remediation, and insufficient resources for curriculum development and teacher training and support.

Drawing from the lessons learned this summer, a number of refinements to the basic model are being made. These include modifications to the remediation curriculum and the development of an expanded life skills curriculum for returning youth.

Impact

At post-test, treatment youth outscored their control group counterparts in both reading and math by approximately one-quarter of a grade equivalent. The impact of participation in STEP varied as a function of sex, race and site.

Treatment girls surpassed control group girls by approximately half a grade equivalent on the reading post-test, while the scores of treatment and control group boys were virtually the same. In math, however, both boys and girls in the treatment group significantly outscored their control group counterparts by one-quarter of a grade equivalent.

Participation in STEP was of greatest benefit to Hispanic youth, whose high school dropout rates are dramatically higher than those for both blacks and whites. Hispanic youth in the treatment group outscored Hispanics in the control group by half a grade equivalent or more in both reading and math. Asians and blacks derived solid benefits as well -- black treatment youth outscored their control group counterparts by one-quarter of a grade equivalent in math. In reading, Asian participants outscored Asians in the control group by half a grade equivalent.

The impact of participation in STEP varied among the sites as well. Treatment youth in Boston scored significantly higher than control group youth in both reading and math. Treatment youth in Fresno and Seattle also scored significantly higher than their control group counterparts in math and reading respectively. However, treatment youth in Portland -- a site that experienced considerable difficulty implementing its remediation component -- were significantly outscored by the control group.

Since the positive impact on reading and math scores is dependent upon control group losses, a thorough exploration of its validity was undertaken. These efforts included the development of a system of incentives to increase control group motivation to perform well, and an analytic investigation of the pattern of pre-test to post-test responses of both treatment and control group youth. An analysis of these efforts generally supports the notion that the control group losses reflect true declines in academic skills.

These academic results, while encouraging, should be improved upon. The positive impact on academic performance was attained by stemming learning losses that would have occurred in the absence of the treatment, rather than by producing actual gains over the summer for treatment youth. For this reason, a number of modifications of the remediation component will be made for next summer. They include: development of curricular modules that will be uniformly adopted by the sites, curriculum content that will link all the program components, and an earlier development process.

Participation in STEP significantly increased fertility-related knowledge for every site, race and sex subgroup. Moreover, there is reason to believe that participation in STEP is beginning to have an impact upon behavior. More treatment boys than control group boys reported abstaining from sex during the program period. Moreover, there are reasons to believe that these differences in the self-reported sexual behavior of treatment and control group boys reflect actual differences in behavior. STEP, then, appears to have demonstrated a positive impact on the sexual behavior of boys, an effect unusual among pregnancy prevention programs. This appears to be a direct result of designing the life skills curriculum to provide a balanced intervention for both boys and girls.

In summary, STEP yielded an interesting and complex set of findings during its first summer, on balance quite encouraging. These results from the first eight weeks of a 15-month treatment demonstrate the potential of this program model as a vehicle for improving the life chances of low-income youth at risk of dropping out of high school. After the demonstration's three-year operational phase, participants will be followed until six months beyond their scheduled graduation dates. During this period, they will be interviewed periodically and their school records will be examined in order to measure school retention rates, graduation rates, pregnancy rates and early labor market experiences.

I. INTRODUCTION

The Summer Training and Education Program (STEP) is an intervention aimed at improving school retention rates and, ultimately, labor market outcomes for poor, educationally disadvantaged youth ages 14 and 15. To accomplish these goals, STEP addresses two major predictors of dropout behavior -- poor academic performance and teenage parenthood.

Following a successful pilot in three cities during the summer of 1984, full-scale implementation of STEP was launched during the 1985 summer. STEP operated in five cities during its first year -- Boston, Massachusetts (one of the pilot sites); Fresno, California; Portland, Oregon; San Diego, California; and Seattle, Washington. These STEP programs were operated by employment and training agencies and local education agencies (LEAs) at each location. Overall development and management of STEP was carried out by Public/Private Ventures (P/PV) of Philadelphia, a national, not-for-profit corporation that seeks effective approaches to job training and education for disadvantaged persons. Funding from the Job Training Partnership Act (JTPA) through the Summer Youth Employment and Training Program (SYETP) and the Ford Foundation represented the primary sources of financial support for STEP. Other public and private funding was also utilized.

This report provides detailed information on the first summer of the STEP demonstration. It covers STEP's implementation in the five sites, the participants served, and the short-term impacts achieved over the summer. The following sections of this chapter provide a framework for these discussions by offering some background on the problem, and full descriptions of the STEP model and the sites.

BACKGROUND

Each year, the number of persons entering the labor market increases. Included among these labor market entrants are large numbers of youth entering for the first time, with no prior work experience. A major subset of this group is comprised of youth without high school diplomas. Their lack of basic skills presents serious impediments to their success in the labor market. Their long-term employment and earning prospects are also seriously curtailed by these deficiencies.

Research has consistently demonstrated that dropping out of school tends to reduce subsequent labor force participation. According to a recent survey of employers (Malizio and Whitney, 1984), only 18 percent of all jobs are open to persons without a high school diploma. Further, the earnings gap between young dropouts and high school graduates increased by 20 percent from the mid-1960s to 1979 (Sum, 1983).

Other research (Levin, 1985) states that the growing number of these unprepared youth in the labor market presents serious consequences that will approach crisis proportions if services are not provided to them. Among the consequences noted by Levin are: "reduced economic competitiveness of the nation as well as states and industries that are most heavily impacted by these populations; massive disruption in higher education; and ultimately, the emergence of a dual society with a large and poorly educated underclass."

As mentioned earlier, the STEP intervention is focused on two major predictors of dropout behavior. First, negative academic experiences -- poor grades and being held back -- are key factors influencing dropout behavior. Furthermore, studies have shown that dropouts are likely to have poor grades in reading and mathematics (Turnage, 1982). One remedy is found in several bodies of research. A study on Job Corps participants and disadvantaged high school and college students has shown that well-designed remedial education programs are capable of producing modest but long-lasting academic gains (Taggart, 1981; Kapsis, 1983). Such modest gains are also evident in a study of black summer school students in Atlanta (see Heyns 1978). Gains in mean grade equivalents on the word knowledge test from the Metropolitan Achievement Test (MAT) were found to be approximately one month per month of instruction or 1.5 months for a six-week program. The study also showed that gains were strongly related to the family income levels of participants, with enrollees from the poorest families achieving smaller gains.

In addition, according to several studies, summer is a time when learning rates for many youth slow, stop or decay. This pattern is most evident in the performance of disadvantaged youth. One researcher (Heyns, 1978) suggests that, for disadvantaged youth, actual learning regression takes place during the summer. While many schools offer remedial education during summer school, the opportunity for a disadvantaged youth to earn wages at a summer job is often a compelling alternative.

The second predictor of dropout behavior for disadvantaged youth being addressed by STEP is teenage parenthood. Teenage mothers and fathers are both more likely to drop out of school before gaining a high school diploma than are youth who postpone parenthood. Recent longitudinal studies suggest that the age at which girls first give birth is one of the strongest influences on whether they complete schooling. As many as three-quarters of female dropouts cite pregnancy and/or marriage as their reason for leaving school (Coombs and Cooley, 1968; Presser, 1977). On average, teenage mothers end up with two years less schooling than classmates who delay childbearing. Adolescent fathers are also affected by early parenthood. One study (Card and Wise, 1978) found that boys who become fathers before age 18 are 35 percent more likely to drop out of school than those who become fathers when they are older.

In addition, research from a variety of sources indicates that work experience during high school pays off in higher rates of post-high school employment, wages and earnings, but that the payoff is proportional to the amount of a participant's educational experience. In other words, high school work experience is more beneficial for a youngster who graduates from high school than for one who drops out (Farkas, 1983). The evaluation of the Youth Entitlement Incentive Projects (Farkas, 1982) has also found that at-risk youth are most responsive to educational interventions tied to opportunities to earn income and acquire job-related skills.

To address the problem of dropout behavior, then, the focus of any prevention program would be well advised to include strategies to improve academic performance and decrease the incidence of adolescent parenthood while not depriving disadvantaged youth of the opportunity to gain work experience and earn summer wages. Predicated on the research cited above, the STEP model does this by offering an integrated program of work experience, academic remediation and instruction in responsible social and sexual behavior.

The findings from the 1984 pilot of STEP lend additional support to the earlier research. Conducted in Boston, Massachusetts; Pinellas County, Florida; and Baltimore, Maryland, the findings show that participation in STEP resulted in significant gains in basic skills, as measured by scores on the Metropolitan Achievement Test (MAT). (While three cities participated in the pilot, the Baltimore site did not serve a control group and was, therefore, not a full pilot site. Despite this fact, Baltimore provided an additional opportunity to observe other elements of STEP in operation.)

Considering the net impact of treatment gains and control group losses, STEP participants in the pilot increased their basic skills substantially beyond what would be expected in the absence of the program. Treatment youth increased their reading levels by just under one full grade equivalent in Pinellas County, and more than one and one-half grade equivalents in Boston. In math, students gained half a grade in Pinellas County and more than one grade in Boston. Treatment youth also achieved improvements associated with STEP's life skills component. At the end of the summer, participants had significantly greater knowledge of methods of birth control than did members of the control group. They also had a better appreciation of the consequences of adolescent parenthood than did controls.

THE STEP MODEL

STEP operates within an already existing program that serves large numbers of disadvantaged youth during the summer months -- the Summer Youth Employment and Training Program (SYETP). SYETP

provides summer work experience and income for youth who would otherwise be unlikely to find jobs. Although SYETP has served many youth over the years, it has realized only limited success in helping these youth make successful transitions to the world of work. STEP demonstrates some additional possibilities for this national program.

Participants in STEP spend two consecutive summers in a program that provides both academic remediation and instruction in life skills in addition to the work experience traditionally available to them through SYETP. In the intervening school year, they take part in a voluntary support component designed to reinforce gains achieved during the first summer and to create expectations for additional gains in the future.

In the short term, STEP focuses on attaining objectives that can be observed within the course of a single summer. Among them are:

- increased achievement in reading and math or reduced summer learning losses in these areas; and
- increased ability to make responsible decisions about sexual behavior.

The central objective is to provide participants with the means to stay in high school and earn a diploma. Offering youth remediation and life skills instruction, while safeguarding their ability to earn wages during the summer, is designed to result in a reduction of school dropout rates due to either poor achievement levels or parenthood. The increased motivation and sense of self-esteem that result from the academic successes of the summer are expected to increase participants' educational aspirations and encourage them to continue their pursuit of a high school diploma.

Four separate but related components comprise the STEP treatment -- remediation, work experience, life skills and opportunities, and school year support.

The remediation component provides at least 90 hours of basic skills instruction, primarily in reading and math, with the aim of improving success in school by producing learning gains or, at least, decreasing the losses usually experienced over the summer months. The STEP model calls for a remediation curriculum that is individualized, self-paced and competency-based. About one-fourth of remediation hours are devoted to computer-assisted instruction (CAI), which is particularly well-suited to competency-based learning and highly motivating for youth without regular access to computers. P/PV developed a set of standards and procedures for the delivery of remediation during the 1985 summer. A flexible system which allowed for local variation, the curriculum guidelines mandated clearly stated competencies, a

comprehensive diagnostic test for reading and math, and a variety of instructional strategies, learning materials and assessment exercises that could be explicitly related to the full menu of competencies.

The life skills and opportunities component presents a curriculum aimed at developing responsible social and sexual attitudes and behavior among participating youth. This component covers such topics as personal decision-making, job equality issues, the consequences of teenage parenting and a variety of ways to avoid pregnancy, including abstinence. The curriculum runs approximately 18 hours, and uses varied presentation formats including lectures, discussions, field trips, role-playing, films and outside speakers.

Work experience is provided through the federally-sponsored Summer Youth Employment and Training Program (SYETP), which funds entry-level jobs in clerical, maintenance and parks and recreation work. Youth are paid \$3.35 per hour, the minimum wage, for their participation in both work experience and remediation/life skills activities.

The school year support component seeks to ensure that treatment youth maintain gains achieved during the summer, continue academic achievement during the school year and return to STEP the subsequent summer.

Target Population

STEP targets low-income youth, ages 14 and 15, who have significant educational deficiencies. To qualify for participation in STEP, students must meet the eligibility requirements of SYETP, which stipulate that participants must be economically disadvantaged.

Fourteen- and 15-year-olds were chosen as the STEP target group for several reasons. First, at 14 and 15, students are thought to be more responsive to remedial efforts than they are at later ages. Many are also making the critical transition between junior and senior high school -- precisely when the dropout rate begins to accelerate. At this age, too, students are interested in obtaining jobs, but are too young to compete successfully with older youth. Finally, at 14 and 15, students reach the threshold of risk for becoming teenage parents and are, therefore, most in need of information about family planning. For many youth at this age, it is also possible to influence those who have not already become sexually active to delay such behavior.

The final eligibility requirement was based on educational performance. Youth performing one or more grade levels below their peers as measured by a recently administered standardized test and/or youth with a history of grade retention are eligible for

participation in STEP, providing the other eligibility requirements are satisfied.

Site Selection

Following preliminary analyses of the 1984 pilot data and the decision to go ahead with a full-scale demonstration, P/PV placed announcements in several employment and training periodicals soliciting letters of interest from sites. These letters, which were due at P/PV by January 31, 1985, were required to cover the following topics:

- the overall program design, including designation of the agencies responsible for each component;
- the design of a remediation program;
- the roles and responsibilities of each key actor throughout the operation of the program;
- whether a site wanted to be considered as a large (500 youth in each cohort) or small (300 youth in each cohort) site; and
- program management, including identification of a single point of accountability for the proposed program.

As a preliminary indication of a site's commitment to the project, signatures of the director of the employment and training agency, the superintendent of schools and other key participants in the proposed program were required.

Fifteen letters of interest were received at P/PV by the due date. These letters of interest were reviewed to determine a site's understanding of the STEP demonstration, and their willingness to make the necessary commitments for implementing a program consistent with the model. Sites were considered to display such understanding by agreeing:

- to serve the STEP target group -- JTPA-eligible 14- to 15-year-olds who were educationally deficient;
- to serve a control group of equal number in the regular SYETP program;
- to randomly assign youth to treatment or control groups;
- to ensure school district participation in data collection from the current year to five years following a youth's participation in the demonstration;

- to adhere to the program design that incorporates remediation, life skills, work experience and school year support for treatment youth; and
- to administer pre- and post-program academic tests and P/PV's baseline and follow-up questionnaires to treatment and control youth.

Ten cities passed this initial screening and were visited to discuss in detail their plans to operate a STEP site. In most cases, sites were visited at least twice. During these meetings, the details of how STEP would be integrated into the regular SYETP process and how school data could be accessed were determined. (Sites with centralized computer systems for school data and payroll were preferred.) These meetings provided the basis on which P/PV assessed sites and made final selections. In addition to expressions of interest from these ten applicants, the pilot sites -- Boston, Massachusetts; Pinellas County, Florida; and Baltimore, Maryland -- were also considered. The major criteria used to select sites were:

1. Management Capability: the organization and structure of the lead agency; data collection capability.
2. Operational Feasibility: the experience of the lead agency in similar programs; proposed plan for integration of STEP into existing operations.
3. Agency Linkages: school district commitment to the project; other agency commitments.
4. Program Design: size, structure, and eligibility criteria; the remediation, work experience and life skills components.
5. Resources: local commitments; possible sources of additional funding.
6. Other Criteria: geographic location; racial/ethnic mix of populations served.

By mid-March, six sites had been selected from the total group: Boston, Massachusetts; Broward County, Florida; Fresno, California; Portland, Oregon; San Diego, California; and Seattle, Washington. The Pinellas County, Florida and Baltimore, Maryland pilot sites elected not to continue participation in the full demonstration.

Although geographic distribution of sites was attempted, the number of applications obtained from west coast cities and the strength of those proposals resulted in the selection of four west coast sites. The Broward County site subsequently withdrew from the demonstration prior to program startup.

In April, a local P/PV on-site coordinator was hired to assist each site in developing and operating the STEP program. P/PV's on-site coordinators play no direct supervisory role in the local agencies operating STEP. Rather, the on-site coordinator is a facilitator, program monitor, data collector, liaison between P/PV and the site and technical assistance provider to these agencies.

Agreements with each site were formalized in Memoranda of Agreement between P/PV and key site actors from the JTPA agency, the school district and other participating organizations. Each memorandum covered in detail the roles and responsibilities of each key site party as well as the responsibilities to be assumed by P/PV.

Chapter III describes program implementation in detail. The following section introduces the demonstration sites.

DESCRIPTION OF THE SITES

STEP works through a complex collaborative arrangement that involves P/PV and its on-site coordinator, the local employment and training agencies, the school district and other local organizations. While programs in all sites adhere to the basic design of the STEP model, local circumstances have been considered in the development of each program.

The STEP sites were each unique in their organizational structures, participant populations and other local conditions. One example of these differences is the number of organizations that collaborated to operate STEP in each site. At one end of the spectrum was Boston, in which five different entities collaborated closely on planning, management and operations. At the other were the two northwest sites, in which one agency assumed primary responsibility for all aspects of the demonstration.

In all the sites, the STEP collaborators were for the most part experienced at summer youth jobs programming and delivery and had worked together before in the traditional roles established for that effort. However, STEP required more complex collaborations among participating agencies than they had experienced before. At least one of the agencies in each site had previously given thought to adding a remediation component -- though not a life skills component -- to their summer programs, but the difficulties of obtaining funding and of achieving the required institutional linkages had been found insuperable.

Cooperation between the public schools and employment and training programs was greatly increased for the operation of STEP. The school systems opened their records (with parental consent) for STEP staff to identify students who were educationally at risk, sat down with staff of outside organizations to plan a

curriculum to be used with in-school students and, to an unusual degree, collaborated in a program that highlighted and addressed problems of academic failure.

The organizations involved with the operation of STEP and the major roles played by each are presented in Table 1. As indicated, a lead agency represents the focal point of accountability for STEP in each site. In conjunction with other community-based organizations (CBOs), city or county agencies, and private organizations, the lead agency is responsible for making sure that each local entity carries out its responsibility to operate each of the program components. The involvement of private institutions in the operation of STEP represents a unique public/private partnership and is another important aspect of the STEP program. The site profiles that follow illustrate the differing administrative configurations employed.

Boston

The program for treatment youth in Boston ran for six and one-half weeks -- from July 8 to August 21. Control youth in Boston worked for eight weeks beginning July 8.

Overall responsibility for STEP was held by the two institutions that administer Job Training Partnership Act (JTPA) funds in Boston: the Neighborhood Development and Employment Agency (NDEA) and the Private Industry Council (PIC). They contracted with Action for Boston Community Development (ABCD), the community-based organization that operates Boston's summer youth program, to take applications for STEP, place control group youth in regular summer jobs, assign treatment youth to one of three remediation sites, conduct the life skills courses, provide counseling to participants, monitor all remediation and work sites, and handle payroll functions.

Remedial education in Boston was delivered by three local universities: Boston University, the University of Massachusetts/Harbor Campus and Northeastern University. Control youth were assigned to regular SYETP jobs in their neighborhoods. Treatment youth were assigned to the university closest to their homes for both the classroom and work experience components. Approximately 50 youth were assigned to each university. Treatment youth held jobs as library assistants, food service workers, boathouse attendants and others in various university departments. Classes were held in the morning for three hours a day, with the exception of one computer-assisted instruction (CAI) session on Friday afternoons at Northeastern. All of the sites had separate teachers for math and reading and the students rotated among the classrooms. The staffing pattern at each university consisted of one university coordinator, two ABCD counselors, two reading teachers (one at Northeastern), two math teachers (one at Northeastern) and four teaching assistants.

Table 1
Organizations Operating STEP, by Site

SITE	LEAD AGENCY	SERVICE PROVIDERS BY COMPONENT		
		REMEDIATION	LIFE SKILLS	WORK EXPERIENCE
Boston	Neighborhood Development & Employment Agency (NDEA)	Northeastern Univ. Boston University Univ. of Mass.	Action for Boston Community Develop- ment (ABCD)	Action for Boston Community Develop- ment (ABCD)
Fresno	Fresno Private Industry Council (FPIC)	Fresno Unified School District (FUSD)	Fresno Unified School District (FUSD)	Fresno County Economic Opportu- nities Commission (FCEOC)
Portland	Portland Private Industry Council	Portland Private Industry Council	Portland Private Industry Council	Portland Private Industry Council
San Diego	San Diego Regional Employment & Training Con- sortium (RETC)	San Diego City Schools	San Diego City Schools	San Diego Regional Youth Employment Program (REGY)
Seattle	Department of Human Resources (DHR)	Department of Human Resources (DHR)	Department of Human Resources (DHR)	Department of Human Resources (DHR)

The amount and scheduling of computer-assisted instruction (CAI) differed by university site primarily because of physical and access constraints. At the University of Massachusetts, for example, youth participated in one hour of computer instruction per week. Youth at Northeastern had one and one-half hours of CAI weekly and Boston University had three hours weekly, split evenly between reading and math. CAI classes at all three universities were held in computer labs housed in buildings separate from remediation classes.

Life Skills and Opportunities (LSO) classes in Boston met twice a week during afternoon hours that would normally have been spent at the work sites. Boston made this scheduling variation because of difficulty incorporating the LSO classes in the morning schedule of their six and a half week program.

The six full-time ABCD counselors assigned to the three universities also assumed responsibility for life skills classes. This arrangement was different from the 1984 pilot experience in which three specialized instructors rotated among the campuses to deliver life skills classes. During the pilot, counseling staff noted that these LSO specialists were not familiar with the students and that it was difficult to coordinate the LSO program with other activities on site. This change, then, was in direct response to those concerns and was also a less expensive option for Boston.

Fresno

The Fresno program ran for seven weeks, from June 24 to August 9, for both groups of STEP youth. As lead agency, the Private Industry Council (PIC) was responsible for operational and fiscal management of STEP. Under contract to the PIC, the Economic Opportunity Commission, a community-based organization, was responsible for outreach, recruitment, intake, economic eligibility determination and work site development for treatment and control group youth. The Fresno Unified School District managed the remediation and life skills components. It developed the curriculum, determined academic eligibility for participation and provided teachers, space and equipment under subcontract arrangements with the PIC.

All remediation (and life skills) classes took place at one middle school site. The school was easily accessible to most students by bus or foot and it had ample classroom, library and computer capacity for all members of the treatment group. Classes began at 8 a.m. and ended at 11:30 a.m. daily; a free lunch and milk were provided.

Participants spent 17-1/2 hours each week at the remediation site; 14 of these hours were spent in reading, math and CAI classes and the remaining hours were spent on life skills. Students attended class in each of the three areas every day, in

groups of about 25 for math and CAI and smaller groups for reading. Remediation instructors were assigned individually to sections of math, CAI or small reading groups. Each remediation section had at least one and as many as four aides, so the student/adult ratio never exceeded 13:1.

STEP youth worked in a variety of jobs for the school district, CBOs and other public agencies. The majority of jobs were with the school district and were typical of those in regular summer youth employment programs. Jobs ranged in duration from three to five hours daily for treatments. Control group youth worked from six to eight hours a day. Total hours were standardized for both groups by varying the number of weeks youth worked.

Life skills classes in Fresno met twice weekly. Two periods of one and one-half hours each were offered to three sections of students. The curriculum was delivered by three life skills instructors. (Four life skills instructors were hired; one, the substitute, did not get an opportunity to teach during the summer.) The life skills instructors in Fresno all had backgrounds relevant to the curriculum content. The teacher/student ratio in life skills classes was 1:25.

Fresno incorporated into STEP a variety of incentives throughout the seven-week program and made a strong push for parent involvement. Proficiency certificates, weekly prizes for perfect attendance, rewards to those who did optional math homework, and "students of the week" named each Friday were among the incentives used. These were subjective selections, made on the basis of cooperation, attentiveness and/or academic achievement. Teachers hung large banners and printed slogans, along with the names of the students of the week, in each classroom. In July, a free parent/student dinner was held at which reading and math proficiency awards were given, enabling parents to see their children rewarded for positive efforts. Students also staged a talent show.

Portland

STEP operated in Portland for seven weeks between June 24 and August 9. The Portland PIC, which administers JTPA funds, assumed responsibility for all SYETP programs in 1984 and is responsible for all aspects of STEP, including school year support activities. This level of responsibility for SYETP was new for the PIC which was involved in a major reorganization effort during the 1985 summer. City-wide consolidation of SYETP, combining JTPA and city resources, was the focus of the reorganization. STEP, therefore, presented yet another wrinkle in an already atypical program for the Portland site. Some organizational problems resulted and influenced the smooth operation of STEP.

The PIC hired its own staff to manage and operate the STEP program. Remediation teachers and aides were hired by the Portland

Public Schools, but were supervised by a full-time STEP coordinator under the aegis of the PIC. STEP staff, with assistance from other SYETP staff and key personnel from the public schools, were responsible for determining program eligibility, handling recruitment and enrollment, and managing implementation of the remediation, life skills and work experience components. STEP staff also handled all payroll functions. The public schools provided planning assistance, classrooms and materials, computer hardware and software, and teachers. Teachers were selected jointly from the regular summer school applicant pool by the PIC's STEP staff and school district staff.

Remediation occurred at a single school site. Students were divided into six classes of 25 students each. Classes met for three and a half hours each morning, with a ten minute break. Each class was to be staffed by three adults: a teacher, an instructional aide and a STEP aide, providing a ratio of one adult to eight students to permit small group work and individualized attention. However, STEP aides were sometimes pulled away for other chores (one of the organizational problems of the new structure), and ratios were often higher.

Half of each morning was devoted to reading and the balance to mathematics. At any given time, instruction occurred in three formats: small group work, when students addressed a similar set of competencies; independent work with the computer; or individualized work with written materials. Teachers and aides worked with small groups or individuals as appropriate.

Treatment youth in Portland received two 90-minute life skills sessions a week. Life skills in Portland was organized differently than in most other STEP sites. Portland's model utilized a single life skills specialist to deliver much of the life skills instruction, including the more technical areas of the curriculum, and regular remediation teachers to deliver designated smaller portions.

Work sites were generally near the students' own neighborhoods. Usually no more than three or four students were placed at each site, all in traditional SYETP public sector jobs. Treatment youth worked three hours each afternoon. Hours for control youth varied widely.

San Diego

San Diego's program ran for seven weeks, from June 24 to August 9. Remediation and life skills classes met five mornings a week for treatment youth and all youth worked four days each week.

A quasi-public agency -- the Regional Employment and Training Consortium (RETC) which includes the PIC -- is responsible for administering all JTPA funds and was primarily responsible for all aspects of STEP. RETC was also responsible for developing

and administering the remediation component in consultation with the school system. RETC contracted with the public school system to operate remediation and life skills, to supply teachers and to serve as work sites for STEP youth. It contracted with the Regional Youth Employment Program, a city agency, to perform recruitment, outreach, intake and eligibility determination, to operate the work experience component and to perform the payroll function. In addition, the school district assisted with recruitment and outreach, screened JTPA-eligible youth for academic eligibility, obtained parental and youth consent forms, identified work sites and supplied participant data, classrooms and computers.

STEP remediation in San Diego took place in five different high schools scattered across a large geographic area. Each school had two classes operating concurrently -- one in reading and one in math. Classes ran from 8 a.m. to 11:45 a.m. and had 12 to 15 students, one teacher and a half-time aide. At the halfway point of the morning, the students switched from one class to the other. Computer-assisted instruction was integrated with the reading and math classes.

The San Diego City Schools had responsibility for administering the remediation component of STEP. This included the hiring of teachers and provision of space and computers. The curriculum development process, however, was shared by the school system and RETC. A curriculum development specialist from the San Diego School District took the lead in developing the math curriculum. Development of the math curriculum for STEP began in April and required little outside help. Development of the reading curriculum was done by the RETC project manager, with assistance from one of the STEP reading instructors.

Two health and human services teachers from the San Diego city school district were selected to teach life skills. The classes were held at the five schools serving as remediation sites. Each LSO teacher took full responsibility for two sites and shared responsibility for the fifth. Twice a week the LSO instructors took over each class for an hour and a half.

Jobs for STEP participants (treatments and controls) were primarily as custodial aides (about 95 percent) with the San Diego City School District. A few clerical and teacher's aide jobs were also available. Work sites were identified by the school district's Career Development department. Youth were assigned to work sites nearest their homes; there were approximately 57 different job sites. The number of youth working at one site varied from one to twelve. Treatment and control participants were placed at the same work sites. Treatment youth worked three hours in the afternoon and control youth worked seven hours daily. Both groups worked Monday through Thursday.

Seattle

Morning (8 a.m. to 11 a.m.) and afternoon (12:30 p.m. to 3:30 p.m.) remediation and work sessions were utilized in Seattle's program. The program ran for eight weeks, from June 24 to August 16.

Responsibility for the design and implementation of STEP rested with the city's Department of Human Resources (DHR), the agency that controls all social service functions including administration of SYETP. With funding from the Seattle-King County PIC, the SYETP was responsible for STEP staffing, recruitment, outreach, intake, academic remediation, life skills, school year support, record keeping and financial management. Other agencies participating in the demonstration included the public schools (providing access to school records, assistance with recruitment and use of computers); the county PIC (wages for participating youth); and the Washington State Employment Security Department (public relations and private-sector fundraising).

Remediation in Seattle took place at a centrally located high school close to public transportation. Five remediation classes plus life skills instruction were conducted in the morning, and five classes in the afternoon. There were approximately 15 participants per class, staffed by a certified teacher and an aide. The teachers were drawn from Seattle Public Schools and surrounding districts. Five teachers and 10 aides were utilized. The aides were work/study students from Seattle University and older SYETP youth.

Computers were located in a separate computer lab, thus class time was devoted to reading and math using print materials, with classes rotating between computer lab and life skills. The weekly schedule allowed for nine hours of reading and math, three hours of CAI and three hours of life skills instruction. Approximately 25 percent of the remediation hours were spent in the computer lab. Students were assisted by their remediation teacher and an SYETP aide assigned to the computer room. Each class visited the lab four times a week for 45 minutes.

Efforts were made to provide participants with jobs either near their homes or near the remediation site. Youth were given scrip or bus passes for transportation between work, remediation and home until the first full pay period. Thereafter, they were expected to pick up their own transportation costs. Treatment and control group youth were given similar work assignments -- in some work sites, they worked side by side. Each group was provided opportunity for both individual placement and group project experiences. Group projects are special work experiences designed especially for summer youth employment. SYETP develops a request for proposal and distributes it to community agencies and city departments. Proposals received are then rated by a team composed of SYETP staff, a representative from another city de-

partment and a representative from a community agency. Although most of the 14- and 15-year-olds held typical SYETP jobs as clerical, maintenance, recreation or child care aides, a few of the projects and placements were quite interesting. One control youth was placed with Medina Children's Services as a public relations assistant; another control youth worked at the University of Washington Law Library where he learned the filing systems for briefs and law journals; nine control youth were involved with a local theater group and performed on stage. Twenty-one treatment youth were placed in the Sasa Havi Puppet Project where they made puppets and performed at local libraries and community centers.

ORGANIZATION OF THIS REPORT

The foregoing discussion of STEP's background, components, program goals and descriptions of the five demonstration sites provides a framework for the remainder of this report, which covers the first summer of STEP implementation in detail. Chapter II describes the research design employed to analyze the study data presented. Chapter III discusses features of STEP implementation in the five sites. Chapter IV profiles the first cohort of STEP youth -- both treatments and controls. The short-term impact of STEP in its first summer is discussed in Chapter V and Chapter VI summarizes the Year One experience and presents plans for continuing the demonstration.

II. RESEARCH DESIGN

STEP's research component is a central part of the demonstration, since its goal is to assess not only STEP's effects on participants but the model's adaptability in various settings as well. The research design is broad-based and includes among its objectives:

- assessment of the program model's feasibility and refinement of its operational procedures;
- measurement of short-term educational gains, and sexual attitudes and behaviors;
- evaluation of STEP's longer-term impact on high school retention, pregnancy and early labor force experiences.

Research began with the pilot program in 1984 and will continue until 1992, five years after completion of the three-year demonstration phase. The major components of the research design are an implementation analysis and an impact analysis. The implementation analysis will assess the replicability of the program model by describing the processes involved in the planning, coordination and operation of STEP. The impact analysis will examine STEP's short- and long-term effects on academic performance, employment-related attitudes and behavior, the development of responsible social and sexual attitudes and behavior, the reduction of adolescent parenthood, school retention and high school graduation, and early labor force experiences.

IMPLEMENTATION ANALYSIS

The implementation analysis will examine factors that contribute both to an understanding of the demonstration's achievements and to later replication of the STEP model. Among the topics to be covered are:

- development and coordination of institutional arrangements -- key individuals, organizations and agreements;
- factors contributing to design of the remediation, life skills and work experience components as implemented at each site;
- problems encountered and solutions attempted during implementation;
- features or occurrences that affected implementation, both positively and negatively;

- characteristics of program participants -- their demographic, attitudinal and behavioral profiles;
- the precise nature of the treatment -- kinds of jobs, hours of work, amount of compensation, type of remediation, and the content of life skills instruction; and
- participants' responses to key features of the program.

Both quantitative and qualitative data sources are used in the implementation analysis. Qualitative sources include observations of program activities -- remediation classes, worksite sessions and life skills instruction -- as well as interviews with key actors. Quantitative data come from application forms, participants' answers to questionnaires designed to tap responses to key program elements, and program time sheets.

IMPACT ANALYSIS

The impact analysis will determine whether STEP participants have better outcomes than a randomly assigned group of youth who participate only in the Summer Youth Employment and Training Program. Since STEP's ultimate focus is on high school graduation rates and their effect on early labor force participation, the research has both short-term and long-term objectives.

In the short term, STEP participants are expected to:

- experience gains in reading and math, or minimize anticipated summer learning losses; and
- exhibit greater knowledge and use of effective contraceptive practices.

Over the intermediate term, STEP participants are expected to:

- exhibit improved academic performance during the school year;
- exhibit reduced levels of sexual activity or improved contraception use; and
- experience a reduction in year-to-year high school dropout rates.

And in the long term, STEP participants are expected to:

- exhibit an increase in high school graduation rates;

- exhibit a reduction in adolescent parenthood; and
- achieve increased employment and earnings in their early labor force experiences.

In the sections that follow, a description of the research design as it relates to this summer's program is presented.

OVERVIEW OF DESIGN STRATEGIES

Of a total of 1,593 STEP-eligible youth, half were randomly assigned to a treatment that combines work experience, remediation and life skills instruction. The remaining youth were randomly assigned to the control group and spent all of their program hours in the regular Summer Youth Employment Program. Treatment and control group youth were compensated equally at \$3.35 an hour -- minimum wage -- for both work experience and remediation.

Because the chances of any given youth being assigned to the treatment group are no greater than his or her chances of being assigned to the control group, random assignment has the effect of creating two approximately equivalent groups. From a statistical perspective the two groups can be presumed equal at baseline. Thus, any differences observed at follow-up can be safely attributed to the effects of participation in the STEP treatment.

In order to minimize the possibility that control group youth have less motivation to perform well at post-test than treatment youth -- and thereby spuriously elevate the estimate of program impact -- an incentive scheme was devised to provide both groups with an external reason for doing their best. Any youth taking the pre-test received one chance in a drawing for one of 15 fifty dollar prizes awarded at each site. He/she received an additional chance for each correct answer scored on the MAT. Thus, a youth could maximize his/her chances of winning by trying to correctly answer as many questions as possible. A youth responding at random, or drawing designs on the answer sheet, or simply not answering the questions -- as, variously, control group youth are hypothesized to do -- would decrease his/her chances of winning. This incentive -- a \$50 prize for trying to perform as well as possible -- was available to both treatments and controls at both pre-test and post-test. If these incentives have their predicted effects, any control group losses observed should be the result of true losses in academic skills.

To analyze the impact of participation in STEP, multivariate analytic techniques were used to adjust for any measurable differences between treatments and controls that survive random assignment, and to increase the precision of the estimation of program impact. Multivariate analysis -- in most cases, a straightforward, ordinary least squares multiple regression --

permitted examination of treatment impact while holding constant the effects of important baseline and demographic characteristics. It was thus possible to detect effects that would not be possible with a simple examination of means or percentages at follow-up. Details of the analytic strategy are discussed in Chapter V.

Sample Development

Random assignment, though the most critical element in the research design, was only one in a fairly complex set of procedures for contact with the sample. Procedures for outreach, application, eligibility verification, intake, random assignment, testing and other data collection efforts are also integral elements in the research design. The following discussion extracts key points of contact with sample members that are common to all sites; Chapter III describes them in greater detail.

Recruitment for STEP was accomplished through the normal outreach and intake procedures already in existence for the regular summer youth program. Youth completed applications, which were checked for conformance with JTPA eligibility requirements regarding age, residence and economic disadvantage. The schools also played a critical role, providing test scores of 14- and 15-year-olds who were potentially eligible. Youth meeting P/PV's educational deficits criteria were then asked to obtain parental consent for participation in a study involving random assignment and, if assigned to the treatment group, for participation in a program involving academic remediation and life skills instruction. Finally, parental consent was sought for the release of school records data for research purposes.

Once randomly assigned, both treatments and controls were administered P/PV's baseline questionnaire, and were pre-tested for reading and math abilities. Several make-up sessions were arranged to reach students who missed the first administration.

The final research contact with members of the STEP sample came at the program's conclusion when participants were tested again for reading and math and were administered a follow-up questionnaire. As with the pre-test session, efforts were made to post-test all members of the assignment sample -- including no-shows, dropouts and absentees. These efforts included scheduling the initial testing session during regular (paid) work or remediation sessions, telling youth that they could pick up their final paycheck after taking the post-test, scheduling make-up sessions for which students would be paid, and telephoning no-shows and dropouts.

DATA COLLECTION STRATEGY

Data on the STEP sample were assembled from a number of different sources and at various points throughout the course of the summer (see Table 2). Among these sources are participants' responses to self-administered questionnaires and interviews, program records, school records and performance on academic tests. These are described below.

Questionnaires

A self-administered questionnaire was used to collect data on demographic characteristics and measures of attitudes, knowledge and behavior relevant to the program, including: educational aspirations, attitudes toward the work roles of men and women, birth control knowledge and awareness of the consequences of adolescent pregnancy. The questionnaire was administered to both treatment and control group youth at the beginning and end of the summer.

Academic Testing

The primary measure of academic performance is the Metropolitan Achievement Test (MAT), a test of achievement in important areas of the school curriculum. The Survey Battery, a subset of the larger Instructional Test, provides a global evaluation of performance in reading, mathematics and several other content areas. It consists of 55 items in the reading comprehension subtest, meant to be completed in 35 minutes, and 40 items in the mathematics subtest, which has a 50-minute time limit.

A number of overlapping batteries in the survey cover the grade range from the beginning of kindergarten to the end of Grade 12. The battery chosen for the STEP demonstration is the Intermediate Level, which is most sensitive for students performing in the 5.0 to 6.9 grade level range. The two forms of this test -- Forms JS and KS -- are functionally equivalent. That is, although the items are different, they are judged to be of approximately equal difficulty. This makes it possible to administer one form as a pre-test and then, ensuring against memory effects, to administer the alternate form as a post-test. This in fact was the procedure in the STEP demonstration.

Program Records

Program records also provide important data about participants and their program experiences. First, application forms for the federal summer jobs program verify income eligibility. Program timesheets record hours of remediation and work experience. Finally, program records are the source of termination data and participants' reasons for leaving.

Table 2

STEP Data Base, by Source

VARIABLES	DATA SOURCE
<u>Demographic Characteristics</u>	
Age	Application Forms
Sex	School Records
Race/Ethnicity	Baseline Questionnaire
Household Composition	
AFDC Receipt	
Single Parent Household	
Public Assistance Receipt	
Language Spoken at Home	
<u>Academic Performance</u>	
Reading	Metropolitan Achievement
Mathematics	Test (MAT) Survey Battery, Intermediate Level I, Forms JS and KS
<u>Education</u>	
Highest Grade Completed	Application Forms
Ever Repeated a Grade	School Records
Grade Repeated	Baseline Questionnaire
Number of Days Absent in Previous Year	
School of Attendance at Intake	
Grade in School at Intake	
Enrolled or Not	
Transferred from Previous Year's School	
School Transferred to	
Attendance (total excused and unexcused absences)	
Grade Point Average	
Passed or Failed	
Date of Graduation	
Date Officially Withdrawn	
Reason for Withdrawal	
Educational Aspirations	
<u>Fertility/Family Planning</u>	
Percent Sexually Active	Baseline Questionnaire
Contraceptive Utilization	Follow-up Questionnaire
Percent with Own Children	Interviews
Knowledge of Consequences of Adolescent Pregnancy	
Number of Pregnancies	
Number of Children Fathered	
Number of Births	
<u>Employment</u>	
Previous Employment	Baseline Questionnaire
Career Aspirations	Follow-up Questionnaire
Attitudes Toward the Work Roles of Men and Women	Interviews
Employment Earnings	
<u>Response to STEP</u>	
Component Most Valued	Follow-up Questionnaire
Component Least Valued	Program Records
Preference for Remediation/ Work Experience Balance	
Response to Computer- Assisted Instruction	
Reaction to Life Skills	
Hours in Work Experience and Remediation	
Termination Date	
Reasons for Termination	

School Records

Since many STEP research hypotheses involve the long-term effects of program participation on school retention and high school graduation, school records are an integral part of STEP's data collection strategy. Participating school districts agreed to provide a substantial amount of school data for each youth in the STEP sample. The student data will include standardized test scores, attendance and grade progression. In addition, the schools have agreed to provide data on local policies and procedures to help interpret student data, and to lend assistance in tracing and gaining access to the records of students who transfer to other school districts.

III. THE IMPLEMENTATION OF STEP

The structures, processes and procedures instituted by the five demonstration sites to operate their STEP projects are discussed in this chapter. The discussion focuses first on the participant selection process, followed by a detailed account of the implementation of STEP's three summer components -- remediation, life skills and opportunities, and work experience -- in the five sites. The discussion highlights similarities and differences in each site's experiences with the design and delivery of these components as well as features that may have influenced program successes and problems.

THE PARTICIPANT SELECTION PROCESS

Outreach and Recruitment

Recruitment and enrollment of youth for participation in STEP were handled through the regular Summer Youth Employment and Training Program (SYETP) in all five sites. Outreach and recruitment of youth started at varying times at each site, but generally began in April. Sites used a variety of strategies to recruit youth for STEP. In general, the strategies employed in the early stages included:

- mailing of STEP information packets to individual schools serving 14- and 15-year-olds;
- newspaper, radio and television announcements; and
- mailing of STEP information packets to community-based organizations.

Later, as it became obvious that a sufficient number of youth were not being enrolled, sites intensified their outreach methods to include:

- visiting schools to talk to youth about STEP;
- handing out STEP information and applications at school buses and during lunch breaks;
- increasing media coverage; and
- expanding the educational deficiency criteria.

In each site, special arrangements were made to steer 14- and 15-year-old SYETP applicants to the potential STEP applicant pool. For many sites, this represented a new focus of service because although JTPA regulations allow for the provision of

service to 14- and 15-year-olds, many programs had given priority to serving older youth. Regular SYETP staff were assigned to work on STEP recruitment exclusively in some sites, while in others they were responsible for recruitment for both STEP and the mainstream program.

Although the SYETP agency was the primary entity involved in enrolling youth for STEP, the role of the public schools was critical. Schools were responsible for providing academic scores of 14- and 15-year-old students to identify potentially eligible youth. This was done in one of two general ways at each site. First, school district lists of all 14- and 15-year-olds who met the academic criteria and their accompanying scores were provided. This method was further fine-tuned in some sites to improve the likelihood that youth identified as academically eligible might also meet JTPA economic eligibility standards. In these cases, participation in the free school lunch program was used as a proxy for economic status. Alternatively, once youth who applied for SYETP were determined age- and JTPA-eligible, school districts provided a list of their scores to SYETP.

Each of these systems had obvious benefits and drawbacks. In the first case, a major drawback was the lack of a precise measure of economic eligibility. However, this was a useful system for identifying youth for whom specific targeted outreach could be employed. The second system was beneficial in that youth were already known to be economically eligible but would not necessarily meet the academic criteria. Obtaining early lists of academically eligible youth from the school district proved to be the most efficient practice and is being instituted in the second year of the demonstration. These early lists allow for early targeted outreach as noted above, but in addition can serve as an early signal regarding the appropriateness of the eligibility criteria being employed and their potential for yielding a sufficient applicant pool. Further, because of the lateness with which these youth tend to apply for summer employment, identification of youth using the school district lists proved more efficient.

The major lesson learned about the outreach and recruitment of youth for STEP is that the effort has to start earlier and be well planned in advance. As the following discussion on eligibility criteria reveals, sites were generally off-target in their initial estimates of the number of 14- and 15-year-old youth that would be yielded by their specific criteria. All sites were optimistic in the early planning stages that youth fitting STEP profiles would be applying for SYETP jobs in large enough numbers that application of the academic criteria would still net the required number. Youth, however, did not apply in the numbers projected by SYETP program operators; thus, the need for better outreach and recruitment strategies. Additionally, we learned that the importance of good collaboration between schools and JTPA agencies in this effort cannot be overstated.

Eligibility Criteria

Fourteen- and 15-year-old youth eligible for SYETP and determined to be at risk of dropping out of school were recruited and selected to participate in STEP. All three criteria (age, SYETP-eligibility and academic risk) are specific and are intended to assure that youth in all sites have the same general profiles. Furthermore, these criteria were carefully selected, based on research findings indicating that youth of this age who have poor educational profiles would benefit most from an intervention like STEP.

Although specific in nature, STEP's eligibility criteria offer sites some flexibility in execution. The education criteria allow sites to select youth on the basis of their scores in reading and/or math and also allow the use of grade retention as an additional or simultaneous screen. All youth must perform at a minimum fourth grade reading level. It should be noted that the guidelines originally set by P/PV during the pilot specified only reading as an appropriate subject screen for determining academic eligibility. However, the reading screen yielded an insufficient number of youth and it was necessary to expand the screens to include math in all sites and, ultimately, language arts in Fresno. Although lists of academically eligible youth produced numbers far in excess of STEP's requirements when the criteria were expanded, the match between academic and JTPA eligibility was more difficult to achieve.

Sites utilized different tests to screen youth for eligibility. In several cases, more than one test was used by a single site. This was due to variation in standardized test administration practices. Some schools within districts utilized different tests, tested in alternate years only, tested only a sample of youth, etc. In four of the five sites it was necessary to modify academic criteria that yielded an insufficient number of eligible youth. Seattle was the only exception. This was probably due to the fact that Seattle started outreach and recruitment activities earlier than other sites and was the only site opting to use the grade retention criterion as an additional screen for eligibility.

Modifications to originally-set academic criteria went beyond the expansion of subject areas and varied by site. Initially, Boston intended to serve only 8th graders who were making the transition to high school. But it became necessary to expand the pool of eligibles by allowing 7th and 9th graders to apply. As noted earlier, language arts was added to the academic screen in Fresno. Portland's initially rigid academic screen, which required youth to perform one to four grades below their current grade in both reading and math, was relaxed to include youth performing at these levels in reading or math. Finally, San Diego expanded its pool by accepting youth scoring one to four levels below current grade rather than one to three as originally planned.

The age criterion offered some flexibility by allowing sites a choice of focusing services on youth of either or both ages. The JTPA eligibility criterion was the only one that offered no room for flexibility.

As the data in the chapter on participant characteristics will indicate, all five demonstration sites succeeded in enrolling youth who met the criteria and who were in need of an intervention like STEP.

Random Assignment

As discussed in the chapter on research design, random assignment of youth to STEP treatment and control groups was a requirement for site participation in the demonstration. Random assignment was intended to occur at a single point after the requisite number of youth were accepted into STEP. However, the irregular patterns of application by youth would have caused a delay in informing them of their acceptance. Therefore, random assignment occurred in rounds. Standard specifications for random assignment were, however, left intact.

The random assignment guidelines were followed closely by all sites with one exception. In Portland, instead of following the prescribed procedure, a method of alternation -- assigning "every other" applicant to the treatment group -- was used. The decision to follow this alternation procedure was made in response to the trickle of applicants into the Portland program but was not approved in advance by P/PV. No significant differences between treatments and controls on key variables were found.

Testing

Youth who participated in STEP were given tests at the beginning and end of the summer to measure their academic achievement levels. The Intermediate level of the Metropolitan Achievement Test (MAT) Survey Battery was used. Sites used either form JS or KS of the MAT at pre-test and the alternate form at post-test. The use of alternate forms at pre-test and post-test ensures against memory effects.

Along with the MAT, questionnaires were administered at both pre- and post-test. They collected a variety of data on STEP youth including background, family composition, educational history and goals, work history, birth control knowledge and use, and other fertility-related issues. Testing generally occurred on the first day of the program for both treatments and controls. In Portland and Boston, however, youth were tested shortly before program start-up.

It was necessary to administer make-up tests in all sites. Make-up testing occurred from three to as many as five days into the

program. Post-testing (including make-ups) occurred during the last week of the program. Table 3 shows the number of youth pre- and post-tested at each site. Sites were successful in testing a high proportion of youth: 94 percent of those randomized took the pre-test and 91 percent of those pre-tested took the post-test. These testing levels represent a considerable improvement over the pilot and were the result of early plans and preparation regarding methods to motivate youth to participate in the testing activity.

In general, for those sites testing on the first day of the program, treatment youth were tested and oriented at their remediation site. Controls were tested and oriented at a central location. In Fresno, because of the district's desire to administer the questionnaire off school property, the local convention center served as the testing site for both groups. Youth were not told their group assignment before testing to avoid differential dropout among the two groups.

The remainder of this chapter presents details on the implementation of the three summer components of STEP -- remediation, life skills and opportunities, and work experience. The section begins with a discussion of the implementation of the work experience component, the most familiar feature of the demonstration.

THE WORK EXPERIENCE COMPONENT

SYETP staff took primary responsibility for implementing the work experience component of STEP in all five sites. Overall, this component realized a fair degree of success. Work experience was the most familiar element of the STEP treatment, and the SYETP staff's knowledge of job training and placement for disadvantaged youth was key in operationalizing this component. However, placements for STEP were unique in several respects. First, SYETP jobs are typically full-time, while STEP requires the placement of youth in part-time positions. Second, SYETP focuses its placement efforts on older youth, while STEP requires slots for 14- and 15-year-olds.

The work experience programs in all sites were similar in that the majority of jobs for treatment youth were with the agencies delivering remediation. Control jobs were with a variety of public agencies. Many jobs for both treatments and controls were custodial. Despite the similarities, some differences also existed. These are highlighted below.

Program Description

The STEP model stipulates that both treatment and control youth earn comparable wages over the summer. In general, sites enforced this design feature by equalizing the number of hours offered to treatments and controls. That is, for treatments, the

Table 3

Number (Percent) of Youth Pre- and Post-tested

	Boston	Fresno	Portland	San Diego	Seattle	All Sites
Youth Randomized	330	317	310	310	326	1593
Youth Pre-tested	303 (.92) ¹	298 (.94)	289 (.93)	292 (.94)	313 (.96)	1495 (.94)
Youth Post-tested	290 (.96) ²	260 (.87)	258 (.89)	273 (.93)	272 (.87)	1353 (.91)

¹Represents the percentage of youth pre-tested of those randomized.

²Represents the percentage of youth post-tested of those pre-tested.

sum of classroom and work hours offered equaled the total work hours offered to control youth. Two exceptions to this design feature existed -- Boston and San Diego. In Boston, treatment youth participated in a six and one-half week program and were offered a total of 198 hours of participation. Ninety of these hours were allotted to the work experience component. Treatment youth were able to work approximately 80 hours and participated in SYETP orientation and testing activities for the remaining 10 hours. Control youth, on the other hand, participated in SYETP for eight weeks and were offered up to 240 hours of work.

In San Diego, the disparity between treatments and controls was much smaller. Both groups of youth participated for seven weeks but treatments were offered up to 211.5 hours (84 of these were work experience) and controls were offered a maximum of 199.75 hours.

Of the remaining sites, Fresno and Portland had seven-week programs and Seattle's program ran for eight weeks. Youth were offered up to 221 hours in Portland, 240 in Seattle and 253 in Fresno. The work experience time allotment for treatments in these sites was 102 hours, 120 hours and 128 hours for Portland, Seattle and Fresno, respectively.

With the exception of Seattle, all STEP youth worked in the afternoons. In Seattle, half the youth worked mornings and half worked in the afternoons.

Implementation Problems

Sites encountered some problems in finding part-time placements for STEP youth. In general, job placement activities started late in all sites and although employers were supportive of the program, many found it difficult or were reluctant to identify jobs that could be done part-time. Locating jobs for control youth was generally not problematic.

Late startup also made it difficult for sites to match youth with available jobs based on skill or desire. Although it was not expected that a high degree of job matching would be achieved, due to the limited skills of STEP youth and the scope of jobs available, less of this occurred than anticipated.

Job development for STEP was also limited. Only in Boston, where treatment youth worked at university-based jobs, and in Seattle, where some youth from both groups worked on specially developed projects, were special efforts at job development made.

Another problem caused by the late start of job placement and development efforts was that all youth in some sites were not assigned to jobs by program start-up; some were assigned to jobs that had insufficient work to last the entire summer and therefore changed jobs two or three times; and some treatment youth in

Fresno worked beyond the post-test in order to accumulate the same number of program hours as controls.

The problems encountered with the implementation of the work experience component are being addressed by modifications to STEP operations planned for the 1986 summer. These modifications are detailed in Chapter VI. For the work experience component they include earlier planning and development, more integration with the classroom components, and better jobs for returning youth.

THE LIFE SKILLS AND OPPORTUNITIES (LSO) COMPONENT

The life skills and opportunities (LSO) component experienced the highest degree of success of the three summer components in all sites. The self-contained nature of the curriculum, the training offered by its developers and student interest in the subject matter were the major attributes of its success.

The curriculum offered approximately 18 hours of instruction, broken into 90-minute segments covering the following topics:

- Values and How They Limit Opportunities for Success
- Sex Role Stereotypes and How They Affect Goals
- Decision Making
- Dating Relationships
- Sexual Relationships
- Parenthood
- Sexuality Facts and Responsibility
- Making Decisions About Sex -- How to Say No
- Saying "Yes" Responsibly

These topics were presented through a variety of formats including role plays, lectures, films and discussion. A visit to a comprehensive health facility was also a part of the life skills program. Clinic visits, which were optional for each student and which required parental consent, served an informational purpose. Very few youth did not receive parental consent to participate in the visit.

Although a standard curriculum was used across sites, strategies for its implementation varied. Life skills classes met twice weekly at remediation sites in all five STEP locations. In general, classes were held during the time allotted for classroom instruction. In Boston, classes met twice a week during afternoon hours that would normally have been spent at worksites; it was difficult to incorporate life skills classes into the morning schedule of Boston's six and a half week program.

Boston's program also varied from the other STEP sites with respect to the background of instructors used to deliver the curriculum. Instructors with backgrounds in sex education, family planning, health education, biology, home economics and related

areas were used in the other sites. In Boston, counselors from the SYETP administering agency delivered the curriculum and sessions requiring technical knowledge were conducted by experts from the agency's family planning unit.

Portland's program was also unique in that some of the life skills classes were delivered by remediation teachers. A single life skills specialist was hired to plan lessons, provide technical assistance and deliver most of the lessons in the curriculum.

THE REMEDIATION COMPONENT

Remediation is the central element of the STEP treatment, comprising more than half the classroom time offered to treatment youth over the summer. STEP remediation focuses on improving skills in reading and math. The instructional approach was designed to be competency-based, individualized and self-paced, and to include up to 25 percent computer-assisted instruction (CAI). A maximum student/teacher ratio of 15 to 1 was required.

Program Description

The specific curriculum used by each site varied. However, all five sites utilized an adaptation of the Community Organized Remediation Education (CORE) curriculum framework. CORE was developed by P/PV and consultants from The Network, Inc. of Andover, Massachusetts.

The CORE curriculum framework specified 14 competency areas for inclusion, as listed below:

- Reading

- Recognizing Words and Identifying their Meanings
- Reading for Explicit Meaning
- Interpreting What You Read
- Reading for Critical Insight
- Interpreting Illustrations

- Math

- Adding and Subtracting Whole Numbers
- Multiplying and Dividing Whole Numbers
- Adding and Subtracting Fractions and Mixed Numbers
- Multiplying and Dividing Fractions
- Adding and Subtracting Decimals
- Multiplying and Dividing Decimals
- Percentage Skills
- Charts and Graphs Skills
- Measurement Skills

Sites were given a planning framework for developing a curriculum that met local needs while adhering to the requirements of CORE. The items included in the CORE framework included defining local competencies and objectives in reading and math, selecting initial diagnostic/locator tests, deciding on the appropriate instructional strategies, choosing instructional materials and assessment exercises related to the competency areas, preparing forms to document student progress, and pulling together the logistical information to assist in CORE implementation.

In developing the remediation component, P/PV had three basic choices. First, it could prescribe a fixed curriculum that met the aims of the demonstration -- a pre-packaged approach such as the Comprehensive Competencies Program. At the other extreme, sites could be given broad development parameters (as was done in the pilot program) and extensive latitude in the development and implementation of the program. Finally, P/PV could develop general standards that define the specific characteristics of the remediation design (e.g., competencies, computer-assisted instruction, self-pacing and individualization), leaving curriculum development staff in each site to flesh out a program.

The pilot experience pointed conclusively to the third choice; local educators were unlikely to work willingly with an externally imposed program unless they had themselves elected to use that program and cost and logistical problems in setting up CCP prevailed.

Therefore, P/PV designed an approach to remediation that specifically delineated the elements of the remediation program, detailed a development process for local educators and curriculum specialists to follow, and provided technical assistance to sites throughout the late spring. This package, called Community Organized Remedial Education (CORE), assumed that sites, on being chosen for the demonstration, would be able to commit staff resources to the development of the curriculum, working with P/PV staff and consultants.

Local development was, therefore, an integral part of the CORE plan and P/PV assumed that appropriate local resources would be made available. However, this did not turn out to be the case. Without a pre-packaged curriculum or the commitment of local resources to prepare one, several sites approached the summer with only partially developed curricula.

Implementation Problems

Although the CORE manual provided a flexible approach to implementing CORE (while upholding its basic tenets), the sites interpreted it as being more rigid. Teachers hired to implement CORE were, therefore, often frustrated by the curriculum. Moreover, many teachers were unfamiliar and/or uncomfortable with the competency-based, individualized approach to instruction.

Implementation problems were manifested in different ways across sites, and the severity of these problems varied. Nevertheless, some problems were generic as described below.

The approach and content of CORE: The expectation by some sites that a fully-developed curriculum would be delivered to them triggered many of the CORE implementation problems. Mixed signals regarding responsibility for CORE development and the extent of allowable local discretion were apparently to blame. On the one hand, while the CORE planning manual stated that sites would develop their own curriculum, it also prescribed the competency areas to be used, class size and an overall learning environment. Confusion over which parts of CORE were required for inclusion and which parts could be altered to fit local needs therefore resulted. Furthermore, the inclusion of a recommended curriculum blueprint led some sites to expect that it would provide more than was intended and subsequently delivered.

Planning and development time: The late selection of demonstration sites -- mid- to late March -- and late submission of the CORE manual to sites were key contributors to some of the timing problems encountered. With the exception of Fresno, all curriculum planning and development started later than it should have. The timing problem was intensified in San Diego and Seattle -- the two sites that intended to use an alternative curriculum approach. In general, sites did not begin planning until mid- to late April, which left about six weeks within which to develop the curriculum.

Curriculum staff: Since most sites were not in a position to develop their own curricula, local experts in this area were not identified early. In fact, in one site -- Portland -- experts were never identified and curriculum development was done by other demonstration staff. In other sites, experts were identified for only one subject area or were not useful to the process. Although a remediation consultant was assigned by P/PV to provide technical assistance to each site, this resource was not sufficient to overcome the problems.

Teacher training: Teachers were not hired until close to program start-up, so there was little time to properly communicate the goals of STEP and the curriculum. In some sites, curriculum development continued beyond program start-up. Teacher training generally took place within a week or less prior to the first day of the program. Many teachers were not familiar with competency-based, individualized instruction or were opposed to it.

Teacher preparation time: There was no time allotted for teacher preparation during the summer. The intensity of the remediation program made it impossible for teachers to meet to supplement their on-the-job learning with additional guidance in defining competencies, individualizing curricula, integrating computer work and the like. In addition, heavy recordkeeping demands did

not leave time for preparation in the abbreviated day for which teachers were paid.

Remediation management: As mentioned earlier, Fresno was the only site with an on-site remediation manager throughout the summer. While P/PV strongly encouraged sites to engage a lead teacher or educational coordinator, this was not mandated. In most sites, off-site resources were used. These resources generally had few linkages to the school system delivering the remediation.

Unfortunately, the momentum of summer programming provided little opportunity to institute mid-program remedies. However, outstanding problems are being addressed through major modifications of the remediation component for the second summer. Modifications include more highly structured curricular modules to be delivered by P/PV to sites, curriculum content that is more closely linked to other program components, a lead teacher responsible for remediation management in every site, and earlier teacher training. It should be noted that the basic philosophy of competency-based, self-paced instruction remains the basis for delivering STEP remediation.

Despite the difficulties outlined in the preceding section, sites were able to provide a meaningful remedial experience for participating youth. Structured observations conducted by P/PV's site coordinators in all five sites found youth to be generally engaged in the activity of learning. The self-paced, competency-based, individualized instructional approach elicited positive responses from teachers in a survey conducted at program end. Across the five sites the math remediation curriculum was more successful and engaging, according to teachers, students and observers. Computer-assisted instruction was also well accepted.

Additional testimony to CORE's success comes from the fact that youth receiving remediation exhibited higher scores at post-test than their control group counterparts in all sites. (See the discussion on program impacts in Chapter V.) Also, 37 percent of STEP youth reported that remediation classes were more interesting than regular school while only 19 percent of the youth found it less interesting. Forty-five percent rated STEP remediation of equal interest to "regular school."

SUMMARY

The STEP summer components -- remediation, life skills and opportunities, and work experience -- were implemented with varying success across the five demonstration sites. Some sites were more successful with some components than others; and the areas of success differed from program to program. Some patterns did emerge, however. These patterns are summarized below.

Remediation

In general, the remediation component had the greatest number of implementation difficulties. The implementation of CORE was subject to a variety of problems including the late startup of curriculum planning and development activities, identification and location of adequate resources, and integration of the curriculum content with other program components.

Life Skills and Opportunities

The life skills component was highly successful -- perhaps the most successful of all the components. Only minor modifications will be made to this curriculum. A new curriculum for returning participants has been developed to incorporate the additional areas covered in Year Two.

Work Experience

The success of the work experience component fell somewhere between that of remediation and life skills. It was perhaps the most homogeneous component across sites. Most youth worked about three hours daily at jobs that were primarily custodial. Many of the jobs were found with remediation deliverers.

The identification of part-time job slots for 14- and 15-year-olds proved to be a challenge but was not impossible. As with the remediation component, earlier attention to the development of jobs and better integration with the other program components are required changes for implementation next summer.

IV. STEP PARTICIPANT CHARACTERISTICS

The characteristics of youth in the first cohort of the STEP demonstration are the subject of this chapter. First, the characteristics of youth in terms of program eligibility requirements are discussed. Next, youth are described in terms of broader demographic, educational, employment and fertility-related characteristics. Both the attitudes and behavior patterns of these youth are discussed. Differences within and across sites are emphasized.

The descriptions contained in this chapter are based on the 1495 youth for whom MAT pretest data are available. Across the five sites, this represents 752 treatment youth and 743 controls. The distribution of respondents in each of the five demonstration sites is as follows: 303 in Boston, 298 in Fresno, 289 in Portland, 292 in San Diego and 313 in Seattle.

While several sources of data are available from which to complete participant profiles, only those that focus on the characteristics of youth at baseline are used in this chapter. Other sources are utilized in the chapter on program impact. The sources used in this chapter are as follows:

- Baseline Questionnaires were administered to all youth at the beginning of the summer. These questionnaires, developed by P/PV, collected data on demographic, educational, employment and fertility-related characteristics, attitudes and behavior of youth as they entered the STEP program.
- MAT Pretests. The Survey Battery, Intermediate Level of the Metropolitan Achievement Test (MAT) in reading and math, form JS or KS, was administered to youth at the beginning of the program.
- JTPA Application Forms were collected as part of the regular summer program. Data on JTPA eligibility, public assistance participation and family income, and a verification of basic demographics are among the items included on JTPA applications.

ELIGIBILITY REQUIREMENTS AND CHARACTERISTICS OF THE SAMPLE

Eligibility Requirements

As noted earlier, all youth participating in STEP are required to meet several requirements. First, they have to be eligible according to JTPA regulations for the Summer Youth Employment and Training Program (SYETP). Second, they have to be 14 or 15 years

old on the first day of the program. Finally, youth must be academically deficient and, therefore, deemed at risk of dropping out of high school. To operationalize this third criterion, two measures of academic deficiency were initially set:

- testing one or more years behind national or local norms for reading as measured by a standardized test; and/or
- being held back one or more years (within the past five) for academic or attendance reasons.

As reviewed in the chapter on program implementation, sites encountered problems enrolling the requisite number of youth when utilizing the academic criteria listed above. Therefore, to increase the applicant pool for STEP, the first measure was amended to include testing below local or national norms on a standardized test in reading and/or math. Further refinement to include language arts was necessary in the Fresno site.

The age and economic eligibility requirements were applied evenly by all sites. However, the measure of academic "fit" for STEP was determined differently in each site. This was the result of several variables, including the type of test administered by each school district, the timing of such administrations, and the inclusion or exclusion of grade retention as a criterion of academic eligibility for STEP. Nevertheless, all sites used criteria that incorporated a minimum fourth grade floor in reading for eligibility.

A full discussion of how educational criteria were implemented by each site was provided in Chapter III on program implementation. The discussion that follows focuses on the characteristics of STEP youth -- first in terms of program eligibility characteristics and then in more general demographic, educational and fertility-related terms.

Eligibility Characteristics

Age. STEP sites enrolled youth between 14 and 15 years of age in each site. Except in Boston, where special efforts to enroll exiting eighth graders were made, youth in the full age range were evenly recruited. The mean age of STEP youth was approximately 14.5 across all sites among both treatments and controls. Some slight differences were apparent in terms of the percentage of 14- and 15-year-olds enrolled by each site. However, none of these differences were statistically significant. Approximately one percent of STEP youth fell outside the 14- and 15-year-old range due to errors made during the screening process. Actual distributions of STEP youth by age and other eligibility variables can be found in Table 4.

Table 4

Distribution of STEP Youth, by Program Eligibility Criteria

	PERCENTAGE OR MEAN DISTRIBUTIONS BY STUDENT ELIGIBILITY CRITERIA											
	Boston		Fresno		Portland		San Diego		Seattle		All Sites	
	T	C	T	C	T	C	T	C	T	C	T	C
<u>AGE</u>												
< 14 Years Old	58%	59%	49%	47%	50%	48%	53%	46%	43%	45%	51%	49%
≥ 15 Years Old	42%	41%	51%	53%	50%	52%	47%	54%	57%	55%	49%	51%
<u>ECONOMIC STATUS</u>												
Public Assistance Participation	63%	59%	71%	67%	47%	52%	73%	76%	61%*	51%*	63%	61%
Mean Household Size	5.05	5.26	5.52**	4.96**	5.06	4.94	5.32	5.74	5.18	5.30	5.23	5.24
<u>EDUCATIONAL CHARACTERISTICS AT BASELINE</u>												
Mean MAT Reading Score	730	723	707	717	716	726	728	726	700	699	716	718
Mean MAT Math Score	685	690	690	692	682	693	723	731	658	664	687	694
Grade Equivalent in Reading	6.5	6.1	5.3	5.8	5.7	6.3	6.4	6.3	4.9	4.8	5.7	5.8
Grade Equivalent in Math	6.3	6.4	6.4	6.5	6.2	6.5	7.4	7.7	5.6	5.8	6.3	6.5
Percent Repeating a Grade	46%	41%	26%	34%	22%	19%	23%	22%	23%	26%	28%	28%
Total in Sample	153	150	150	148	146	143	145	147	158	155	752	743

* The difference between treatments and controls is significant at the .10 level.

**The difference between treatments and controls is significant at the .05 level.

Economic Eligibility. All youth in STEP had to meet the income requirements of the Summer Youth Employment and Training Program. These requirements stipulate that youth must be economically disadvantaged as proven by their family income level. Variation in the manner by which income data are collected and the inconsistency of such collection across sites makes it difficult, if not impossible, to analyze this data. However, one indication that STEP youth were indeed economically disadvantaged is the percentage reporting family participation in a public assistance program. Public assistance participation was derived by looking at JTPA application data for reports of AFDC and/or food stamp receipt. Overall, 61 percent of all youth reported such participation. Portland's youth reported the lowest participation rate -- 49 percent overall -- and San Diego showed the highest rate, with 74 percent of all youth there reporting participation. In the aggregate, treatment and control differences were not significant. Within sites, the only significant difference was evident in Seattle, where 61 percent of treatments participated in public assistance programs compared to 51 percent of controls.

Educational Eligibility. As detailed in Chapter III, each site used a different measure to screen youth for participation in STEP. Once youth were determined educationally eligible by locally-applied standards and randomized to treatment or control group conditions, they were required to take math and reading subtests from the Intermediate Level of the Metropolitan Achievement Test (MAT) Survey Battery. The MAT was given in order to provide a uniform basis for verifying that the youth accepted into the program were indeed educationally deficient. Scores resulting from the MAT pretests indicate that sites were successful at enrolling youth in need of academic remediation.

Typically, youth who entered the first grade at age six would be enrolled in the 9th and 10th grades at ages 14 and 15. STEP youth in the first demonstration cohort were, on average across all sites, enrolled in the fifth month of the eighth grade. Furthermore, the MAT scores indicate that youth were performing well below what is expected of youth at their current grade levels. In the aggregate, treatment youth had a mean score of 716 in reading and controls a mean 718. These mean scores translate to 5.7 and 5.8 grade levels respectively. While STEP youth performed better in math, scores were still below current grade levels. The mean score in math for treatments was 687 (6.3 grade level) and for controls it was 694 (6.5 grade level). No significant differences between treatments and controls in academic ability were evident based on these scores. These two measures -- grade in school and MAT scores -- provide verification of the academic deficiencies of STEP youth.

While scores for treatments and controls were similar, more variation is obvious when scores are examined by site, race, age and sex. In reading, youth in Boston and San Diego scored highest, with mean scores of approximately 727 or a grade equivalent of

6.3 (third month of the sixth grade), followed by youth in Portland, Fresno and Seattle respectively. In math, youth in San Diego outscored the second highest site -- Fresno -- by more than one full grade equivalent. The mean San Diego score of 727 placed youth at the 7.5 grade level. Fresno was followed by Boston, Portland and Seattle respectively. The high San Diego math score may reflect the large Asian-American population served in that site. (See Tables 5A and 5B.)

Examination of scores by race shows that for the aggregate sample, whites had the highest reading scores. Furthermore, the reading score of whites was significantly higher than that of other races in Portland and for the aggregate sample. Hispanics had the next highest scores, followed by blacks and Asians respectively. In math, Asian youth scored highest for the aggregate sample and within every site but Seattle. The overall mean score for Asians in math was approximately 742 or a grade equivalent of 8.1. Hispanics followed with a mean score of 706 (6.8 grade equivalent). The difference between the math scores of Asians and those of other racial groups was significant at the .05 level for the aggregate sample and in each site except Seattle. In Seattle, the math scores of Hispanics were significantly higher than the scores of other groups in that site.

With respect to age, no significant differences between the two age groups in the aggregate sample were found for reading. In math, however, 15-year-olds scored significantly higher (at the .10 level) than did 14-year-olds in the sample as a whole.

Finally, females scored higher in both reading and math than did their male counterparts across sites. In reading, females had a mean score of 724 compared to the male score of 711. This difference was significant at the .05 level. In math, the margin of difference was smaller and not significant, with females receiving a mean score of 692 compared to a mean score of 689 for males.

Another measure of academic fit for STEP was grade retention. Seattle was the only site that, in conjunction with test scores, utilized this screen for STEP participants. There was strong variation by site with respect to grade retention. Youth in Portland reported the lowest incidence of grade retention (21 percent) and Boston's youth reported the highest rate (43 percent). Fresno, San Diego and Seattle reported 30 percent, 23 percent and 24 percent grade retention respectively.

Table 5A

Distribution of STEP Youth's MAT Scores
in Reading, by Selected Variables

	MEAN MAT READING SCORE AT BASELINE					
	Boston	Fresno	Portland	San Diego	Seattle	All Sites
<u>AGE</u>						
< 14 Years Old	729	713	719	722	699	717
≥ 15 Years Old	724	712	722	732	700	717
<u>RACE</u>						
Asians	776	N/A	722	731	688	713
Blacks	725	707	715	723	704	715
Hispanics	726	719	718	729	728	723
Whites & Others	729	708	736	737	700	725
			(F=2.65)**			(F=2.56)*
<u>SEX</u>						
Male	723	712	715	722	687	711
Female	730	712	728	731	715	724
			(t=2.13)**		(t=4.42)**	(t=4.20)**

* Significant at .10 level.

** Significant at .05 level.

Table 5B

Distribution of STEP Youth's MAT Scores
in Math, by Selected Variables

	MEAN MAT MATH SCORE AT BASELINE					
	Boston	Fresno	Portland	San Diego	Seattle	All Sites
<u>AGE</u>						
< 14 Years Old	689	680	687	721	656	687
> 15 Years Old	686	701	688	734	664	694
		(t=2.36)**				(t=1.66)*
<u>RACE</u>						
Asians	815	N/A	747	771	704	742
Blacks	684	684	672	697	643	674
Hispanics	694	698	708	736	724	706
Whites & Others	687	695	699	731	644	688
	(F=6.30)**		(F=11.99)**	(F=23.12)**	(F=13.84)**	(F=49.95)**
<u>SEX</u>						
Male	688	695	685	725	659	689
Female	688	686	691	729	662	692

* Significant at the .10 level.

** Significant at the .05 level.

OTHER DEMOGRAPHIC CHARACTERISTICS

Sex. Across all sites, 53 percent of both the treatment and control youth in STEP was male; 47 percent of the sample was female. San Diego was the only site where more females than males were enrolled (52 percent vs. 48 percent). No significant differences between the percentage of males and females served across all sites were found.

Race. A racially diverse group of youth comprised the first cohort of the STEP demonstration. Blacks were the largest group served by STEP -- 59 percent of the total sample. The second largest group served was Hispanics (16 percent), followed by Asians (14 percent), whites (10 percent) and others (1 percent).

There was considerable diversity across sites with respect to the combination of racial groups served. Asian youth were mostly concentrated in the Portland, San Diego and Seattle sites as displayed in Table 6. San Diego's sample was comprised of approximately 30 percent Asians, Seattle's about 27 percent and Portland's approximately 11 percent. Asians in the sample came from varied backgrounds. For the aggregate sample, Vietnamese youth represented the largest Asian constituency (29 percent). This was followed closely by Chinese youth (28 percent), Laotian youth (12 percent), Cambodian youth (11 percent), and Filipino and Hmong youth (7 percent each). Youth from Japanese, Thai and other Asian backgrounds were also served in STEP; however, these youth constituted one percent or less of the sample as a whole.

English Proficiency. In the first summer of the demonstration, STEP served a large number of youth for whom English was not their only language. Across all sites, nine percent of the sample reported limited English proficiency -- that is, they came from homes where some language other than or in addition to English was spoken and they had difficulty reading, speaking or understanding English at times.

The languages most frequently cited in addition to English were Spanish, Vietnamese and Chinese. Consistent with the racial diversity of the participants in each site, Spanish was the second language spoken by youth from Fresno (a primarily Mexican-American group) and Boston (a primarily Puerto Rican group). Asian languages were prominent among youth from San Diego, Portland and Seattle.

Household Size. For the aggregate sample, there were no significant differences in mean household size among treatments and controls. Both groups reported average family sizes of just under 5.25 persons. Fresno was the only site where a significant difference in the household size of treatments and controls was found.

Table 6

Demographic Characteristics of the STEP Sample at Baseline

	PERCENTAGE OR MEAN DISTRIBUTION BY SITE					
	Boston	Fresno	Portland	San Diego	Seattle	All Sites
<u>SEX</u>						
Male	50%	56%	54%	48%	56%	53%
Female	50%	44%	46%	52%	44%	47%
<u>AGE</u>						
< 14 Years Old	58%	48%	49%	50%	44%	50%
≥ 15 Years Old	42%	52%	51%	50%	56%	50%
<u>RACE</u>						
Asians	2%	0	11%	30%	27%	14%
Blacks	69%	51%	64%	49%	62%	59%
Hispanics	13%	43%	3%	17%	2%	16%
Whites & Others	16%	5%	22%	4%	9%	11%
<u>OTHER DEMOGRAPHIC CHARACTERISTICS</u>						
Household Size	5.15	5.24	5.00	5.53	5.24	5.23
Percent From Female-Headed Households	56%	52%	50%	47%	45%	50%
Percent Reporting Limited English Proficiency	4%	3%	8%	16%	13%	9%
Percent With Prior Work Experience	0	22%	7%	0	0	6%

Treatments in Fresno reported a mean family size of 5.52 whereas controls reported a mean of only 4.96.

Family Composition. One-half of the youth in the STEP sample came from families headed by a mother, with no father or step-father present in the household. Fifty-six percent of the youth in the Boston site reported coming from this type of family -- the highest incidence of all sites. The lowest percentage was in Seattle, where 45 percent of youth described their household in this way. No variation between treatments and controls is evident with respect to this variable for the aggregate sample. The only significant difference (at the .10 level) occurred in Portland, where 55 percent of the treatments reported being from "mother only" families compared to 44 percent of the controls.

Employment History. Only 6 percent of STEP youth reported having had previous paid work experiences on either a part- or full-time basis. These youth came from the Fresno and Portland sites only. In the three remaining sites less than one percent of the youth reported any job history. Therefore, for the overwhelming majority of youth, participation in STEP provided an initial introduction to the world of work.

ADDITIONAL EDUCATIONAL CHARACTERISTICS

Highest Grade Completed. As noted earlier, youth in STEP were enrolled in the fifth month of the eighth grade on average. As shown in Table 7, the range of grades represented is quite large. Boston served the highest concentration of 8th graders as intended by their program eligibility requirements. In all other sites, the highest concentration of youth were exiting from the 9th grade (or above). No significant differences based on grades completed were found.

Absenteeism. Another indication of potential dropout behavior often used by researchers is the school absentee rate. Youth in STEP were asked to report the number of days absent in the school year completed just prior to enrolling in STEP. Across the five demonstration sites, 27 percent of STEP youth were absent in excess of 10 days. Fresno had the highest absentee rate of all sites (37 percent), followed by Seattle, Portland, Boston and San Diego, respectively.

Educational Aspirations. The educational aspirations of STEP youth at baseline were fairly consistent for all five sites, with 71 percent of sample youth reporting that they expected to pursue some form of post-high school education. Seventy-eight percent of the youth in San Diego expected to continue education beyond high school -- the highest of all sites -- while only 63 percent of Boston's youth expected to do so. While research generally shows that educationally disadvantaged youth are disenchanted

Table 7

Additional Educational Characteristics of the
STEP Sample at Baseline

	PERCENTAGE OR MEAN DISTRIBUTION BY SITE					
	Boston	Fresno	Portland	San Diego	Seattle	All Sites
<u>Grade in School During SY '85-86</u>						
6th Grade or Less	<1%	0	0	<1%	0	<1%
7th Grade	7%	8%	3%	8%	9%	7%
8th Grade	75%	28%	36%	38%	39%	43%
9th Grade and Above	17%	64%	60%	54%	52%	50%
Percent Absent More than Ten Days During SY '85-86	24%	37%	29%	17%	29%	27%
Percent Expecting to Pursue Post- High School Education	63%	70%	72%	78%	74%	71%

with school and their prospects for educational advancement, STEP youth displayed contradictory attitudes. Intervening in these youth's lives at a time prior to such disenchantment provides a good opportunity for STEP to reinforce the importance of a high school diploma and perhaps reduce the incidence of dropping out.

FERTILITY-RELATED CHARACTERISTICS

To evaluate the life skills and opportunities curriculum, a series of questions related to youth's sexual attitudes and behavior, their attitudes toward the consequences of early parenting and related issues were asked on the baseline questionnaire. Because of the sensitive nature of many of these questions, special steps were taken to ensure the confidentiality of all questionnaire materials and to encourage honest answers by youth. Youth were not required to place their names on the questionnaires and were instructed to enclose the completed questionnaire in a sealed envelope bearing only an ID number. Details on questionnaire and test administration procedures were provided in Chapter II.

To test STEP youth's knowledge of birth control issues, two questions were asked at baseline. The first tested knowledge about the availability, without prescription, of eight birth control methods or devices. The second was a series of questions that tested factual knowledge about the use and effectiveness of different birth control methods. For both sets of questions, youth were given the option to select an answer which indicated their unfamiliarity with the method in question in addition to the true/false options.

Based on the responses to both questions, STEP youth proved to be largely uninformed and/or misinformed on birth control-related issues. On the question testing knowledge of methods available without prescription, the average number of correct responses was approximately three (of eight). Youth indicated that they had never heard of an approximately equal number of methods. Similar results were evident from the responses to the birth control knowledge questions. Here, of the twelve statements provided, an average of just under five (4.84) were correctly answered. On average, four of the methods were unfamiliar to the youth.

These data indicate that STEP youth are in need of and could benefit from the type of instruction offered through the Life Skills and Opportunities curriculum, despite reports by 60 percent that they had taken a previous course in sex education. This is especially true when one examines the percent of youth who said they had already engaged in sexual activity. (The mean age for the onset of sexual activity for the aggregate sample was just under 12.) More than half of both treatment and control youth (51 percent) reported some previous experience in sexual

activity. Youth in the San Diego site reported the lowest incidence of sexual experience -- approximately 38 percent. Portland had the highest incidence, with 60 percent of youth reporting sexual experience. Differences between treatments and controls were found to be significant (at the .05 level) in two sites -- Boston and San Diego. A higher proportion of control youth in Boston reported sexual experience than did treatment youth (65 percent versus 53 percent). In San Diego, 45 percent of the treatment youth compared to 32 percent of the controls reported having had sexual experience.

A lower percentage of STEP youth reported sexual activity in the two months prior to program entry. Approximately 29 percent (28 percent of treatments and 31 percent of controls) said they had engaged in sexual activity during the prior two months. Boston was the only site in which a significant difference between treatment and control youth was found for this variable. Twenty-eight percent of treatments compared to 39 percent of controls reported some sexual activity in the two months prior to program start-up.

Simultaneous with their sexual experiences, these youth exhibited an inconsistent pattern with respect to birth control usage. Only about 40 percent of youth from both randomized groups stated that they used birth control during the last two months of sexual activity. A larger group -- 41 percent of treatments and 46 percent of controls -- reported never having used any method of birth control. These results showed no significant differences among treatments and controls with respect to knowledge and behavior related to fertility issues.

The attitudes of STEP youth with respect to the consequences of teenage pregnancy and responsibility for birth control seem inconsistent with the behavior patterns described above -- that is, they seem aware of the difficulties but are nevertheless careless in their behavior. Seventy-four percent of treatments and 71 percent of controls felt that a teenage girl who becomes pregnant is most likely to drop out of school for a considerable period of time rather than continue attendance on a full- or part-time basis. Related to responsibility for birth control, 71 percent of both treatments and controls felt that the boy and girl should be equally responsible for making sure that some birth control is used.

Twenty-two STEP females reported being currently or previously pregnant. These young women were distributed fairly evenly across the sites. Similarly, 24 STEP males reported having been responsible for a pregnancy or fathering a child. These and other fertility-related data are exhibited in Table 8.

Table 8

Fertility-Related Characteristics of
STEP Youth at Baseline

	PERCENTAGE OR MEAN DISTRIBUTION OF STEP YOUTH AT BASELINE											
	Boston		Fresno		Portland		San Diego		Seattle		All Sites	
	T	C	T	C	T	C	T	C	T	C	T	C
Mean Number Correct Answers on Birth Control Availability Question (Maximum Score Possible = 8)	3.28	3.11	2.65*	3.04	3.12	3.39	2.97	2.99	2.92	2.79	2.99	3.06
Mean Number Correct on General Birth Control Knowledge Questions (Maximum Score Possible = 12)	5.14	5.07	4.73	5.11	4.78	5.04	4.43	4.88	4.69	4.51	4.76	4.92
Percent of Full Sample with Prior Sexual Experience	53%**	65%	52%	49%	59%	60%	45%	32%**	49%	47%	52%	51%
Percent of Full Sample Sexually Active in the Past Two Months	28%	39%	24%	30%	35%	38%	23%	19%	30%	28%	28%	31%
Percent of Sexually Active Sample Using Contraception in the Past Two Months	34%	39%	34%	26%	45%	51%	33%	41%	50%	44%	40%	41%

* Significant at the .10 level.

** Significant at the .05 level.

Table 8 (cont'd)

	PERCENTAGE OR MEAN DISTRIBUTION OF STEP YOUTH AT BASELINE											
	Boston		Fresno		Portland		San Diego		Seattle		All Sites	
	T	C	T	C	T	C	T	C	T	C	T	C
Percent of STEP Males Who Fathered or Were Responsible for a Pregnancy (n)	<1% (1)	<1% (1)	<1% (1)	<1% (1)	3% (4)	3% (5)	<1% (1)	1% (2)	3% (5)	2% (3)	3% (12)	3% (12)
Percent of STEP Females Currently or Previously Pregnant (n)	<1% (1)	2% (3)	<1% (1)	<1% (1)	3% (4)	3% (5)	1% (2)	1% (2)	1% (2)	<1% (1)	3% (10)	3% (12)
Percent Who have Taken a Prior Sex Education Course	45%	50%	56%	47%	74%**	62%	79%	77%	55%	51%	62%	58%
Mean Age at Onset of Sexual Activity	12.14	11.72	11.87	11.88	11.30	11.52	11.79	11.38	11.31	11.68	11.67	11.65
Percent Knowledgeable About the Consequences of Teenage Pregnancy	86%	81%	65%	65%	74%	78%	72%	69%	73%	62%*	74%	71%
Percent Believing that Both Partners Should Be Responsible for Birth Control	70%	63%	63%	70%	76%	80%	77%	69%	69%	73%	71%	71%

* Significant at the .10 level.

** Significant at the .05 level.

SUMMARY

STEP served a racially diverse group of youth who are homogeneous in their need for remedial education, summer employment and instruction in responsible social and sexual behavior. All were JTPA-eligible and 61 percent came from families receiving public assistance. Their MAT scores indicate that STEP youth perform well below their grade levels in reading and math. At an average age of 14.5, many STEP youth are sexually active but lack knowledge about pregnancy prevention strategies and the impact of early parenthood on future goals. These youth tend to come from families with over five members, and fully half of them live in homes with an absent father or stepfather.

This chapter has provided a profile of the youth served by STEP during the demonstration's first summer and, as such, examined only the characteristics of youth at baseline. Chapter V looks at the short-term impacts of STEP from baseline to follow-up.

V. PROGRAM IMPACT

As a demonstration that focuses on high school graduation rates and early labor force participation, STEP has both long-term and short-term objectives. As a long-term proposition, STEP seeks to reduce the high school dropout rate among disadvantaged young people who have records of poor academic performance and who are at risk of adolescent parenthood. It seeks to do so by first attaining several short-term objectives: improving participants' basic skills and attachment to school and increasing their motivation and ability to avoid adolescent pregnancy.

This chapter presents an analysis of the impact of program participation on these short-term objectives -- learning gains, return to school and life skills outcomes. The presentation of program impacts is preceded by two discussions. The first is a presentation of the analytic strategies by which the program estimates were derived. The second is a discussion of the hours of treatment -- remediation and/or work experience -- actually received. This latter discussion is important because it can contribute to a better understanding of the reasons behind the presence or absence of significant program impacts.

BASIC ANALYTIC STRATEGIES

Measurement of the impact of participation in STEP relies heavily upon multivariate analysis -- in most cases, a straightforward ordinary least squares multiple regression -- to statistically adjust treatment effects for measured differences between treatments and controls on important baseline and demographic characteristics. Moreover, by incorporating more information into the estimation process, multivariate analysis takes advantage of interrelationships among variables and so produces greater precision.

An additional advantage of multivariate analysis is that it permits an examination of the impact of the treatment while holding constant the effects of all other variables specified in the model. It is thus possible to determine effects that are not apparent from a simple examination of means or percentages at follow-up. Multivariate analysis can be an asset even in a design that includes random assignment, since differential attrition from the program can undo the effects of the initial assignment. Moreover, random assignment does not guarantee that there will be no significant differences between treatments and controls. Though it has the effect of reducing the likelihood that such differences will exist, often a small number of significant differences will occur in spite of random assignment.

In certain situations, departures from this basic strategy are called for. For instance, the use of a linear model may not be

appropriate when the dependent measure of interest is a dichotomous variable (e.g., whether or not a participant uses contraception during the program period). In such cases, a nonlinear maximum likelihood estimation technique, either logit or probit, is used to estimate treatment effects. In situations where the dependent variables are jointly determined, a two-stage least squares simultaneous estimation procedure is warranted.

In this analysis plan, the effect of treatment is examined in a number of ways. The basic technique is to create a dichotomous variable indicating treatment or control group status. The bulk of the analysis of STEP followed this approach. However, in order to explore the effects of participation in STEP on learning outcomes more fully, the dichotomous variable indicating treatment or control group status was replaced with a variable indicating the intensity of participants' exposure to the treatment -- e.g., the number of remediation hours received. This analysis provided an estimate of the extent to which increasing exposure to the STEP remediation component resulted in improved learning outcomes.

In order to investigate this hypothesis, an analysis of total hours spent in remediation classes was used to explain post-test scores in reading and math. Since it is likely that how well a participant does on the post-test is determined concurrently with how often he or she attended the program -- that is, the same characteristics that predict better attendance are likely to predict greater improvement -- a straightforward regression model with hours of remediation included in the same manner as any other explanatory variable would be inappropriate. In such cases, hours of remediation do not constitute an independent variable, but a dependent variable that is determined simultaneously with post-test scores. Ignoring this relationship would generate biased and inconsistent estimates of the coefficients in the model. The way to correct this problem is to estimate the dependent variables -- in this case, remediation hours and reading and math post-test scores -- simultaneously. A system of three equations (hours, reading and math) was estimated using the two stage least squares (2SLS) technique.

In addition to examining the impact of participation in STEP on the sample as a whole, another aspect of the analytic strategy was to examine the impact on key subgroups within the sample -- the individual sites, racial and ethnic subgroups, and males and females. The intent of this analysis was to find out whether the impact (or lack of impact) applied equally to all subgroups or was particularly effective for certain groups. Some topics (for instance, the interaction between site, race and sex) could not be approached through separate analyses; in such cases, a pooled multiple regression model was employed.

In summary, a variety of analytic strategies were utilized in evaluating the impact of participation in STEP. The fundamental

approach used a dummy variable (indicating treatment or control group status) in an ordinary least squares multiple regression model, controlling for pre-treatment differences in the characteristics of the sample. Additional analyses were performed to deal with situations where the assumptions of the ordinary least squares model were likely to be violated, to handle situations where dependent measures are jointly determined, to refine the measurement of the treatment variable and to examine program impacts for selected subgroups of participants.

An additional issue of concern in the interpretation of impact data is attrition. It is important that the sample generating estimates of program impact be representative of the entire sample prior to significant sample loss. Differential attrition from the treatment and control groups -- as one moves from the number of eligible youth who were randomized to the number who have responses to all the variables in the learning and life skills analyses -- might pose a threat to the internal validity of the study if program outcomes are a function of differences in the characteristics of participants remaining in or exiting from the program. On the other hand, if treatment and control group youth who drop out have characteristics that are similar -- e.g., if the variables for which they have significant differences are relatively few -- then the problem is less serious. Treatment and control group differences aside, it is also important to the generalizability of the results that other key characteristics of the randomization and analysis samples remain comparable. If a disproportionate percentage of one racial group is lost to the analysis sample, for instance, an overestimate or underestimate of the true program impact on STEP-eligible youth could result.

A number of steps were taken to investigate the possibility of attrition bias. First, the characteristics of the samples were compared to determine how representative of the population of youth randomly assigned to STEP the final samples were. The key results of this analysis are presented in Table 9, which depicts selected characteristics of the original randomized sample, the learning analysis sample and the life skills analysis sample. It suggests that the samples are similar with respect to age, race and sex.

Multivariate analytic techniques were employed to determine more definitively whether inclusion in the learning or life skills analysis samples (which require non-missing data on the variables in the appropriate prediction models, as well as pre- and post-test scores on the MAT) varied by age, race, sex or treatment or control group status. The full results of the probit analysis appear in the Appendix. The indication is that, net of background factors, treatment or control group status did not predict whether an individual was likely to be included in either the learning or the life skills analysis sample. Thus, differential attrition from treatment and control group samples does not pose a threat to the interpretation of these data.

Table 9

Selected Characteristics of the Randomization,
Learning Analysis and Life Skills
Analysis Samples

	RANDOMIZATION SAMPLE (n=1593)			LEARNING SAMPLE (n=1144)			LIFE SKILLS SAMPLE (n=937)		
	Treatment	Control	Total	Treatment	Control	Total	Treatment	Control	Total
Male	52.8	53.3	53.0	50.4	51.2	50.8	46.1	50.1	41.8
Female	47.2	46.7	47.0	49.6	48.8	49.2	53.9	49.9	51.9
14 & Under	49.9	49.3	49.6	51.2	48.1	49.6	51.3	46.9	49.1
15 & Under	50.1	50.7	50.4	48.8	51.9	50.4	48.7	53.1	50.9
Black	60.5	56.6	58.6	61.2	56.5	58.8	60.6	57.5	59.0
Hispanic	15.5	16.2	15.9	15.2	15.5	15.4	15.1	14.6	14.9
Asian	13.8	14.3	14.1	15.4	15.0	15.2	16.4	14.2	15.3
White & Other	10.2	12.8	11.4	8.1	13.0	10.6	8.0	13.6	10.8

NOTE: Of the 1,593 youth initially randomized, 270 fell away because they were missing either the pre-test or post-test; the remainder fell away because of non-response on at least one variable in the appropriate equation.

MODELING PROGRAM PARTICIPATION

Understanding program participation -- e.g., knowing what characteristics predict whether a youth will remain in the program or drop out after a relatively short time -- is important from a number of perspectives. First, in order for STEP to have a beneficial effect, participants must receive substantial exposure to the treatment. Second, what is learned from the participation analysis can be used to fine-tune the program design in order to increase the number of youth who complete the program.

The dependent measure for the participation analysis is the amount of time a participating youth spent in program activities. This includes two specific variables: classroom hours in remediation and life skills, and total program hours in both work experience and the classroom.

On average, participants spent a total of 181 hours in the STEP program this summer. In the aggregate -- e.g., considering all sites together -- control group youth averaged 183 hours of work experience. Treatment youth averaged 180 hours, a little more than half of which were classroom hours (94 hours), and the remainder work experience hours. As was discussed in Chapter III, however, there were differences by site in the maximum number of hours made available to treatment and control group youth. Table 10 presents data on the distribution of total program hours by site, expressed as a percentage of the maximum number of hours available. In two of the sites -- San Diego and Fresno -- these data point to sizable differences between treatments and controls in the number of total program hours received. It is somewhat surprising that in both cases, treatment youth participated in a higher percentage of available program hours than did control group youth.

In the aggregate sample, multivariate analysis was used to investigate the influence of a number of background variables on hours of program participation. It showed that there was very little in the way of variation in total program hours for the aggregate sample. Treatments and controls, males and females, 14- and 15-year-olds all participated in the program for about the same number of hours. This analysis does reveal significant race differences: blacks and Asians had more program hours than other groups. Total program hours were also significantly related to baseline reading scores and attendance patterns in regular school.

Table 11 presents similar data for classroom hours for treatment youth. Like total program hours, there was substantial site variation in the number of remediation hours (expressed as a percentage of the maximum number of remediation hours available) that were available to participating youth. Treatment youth in San Diego completed the highest percentage of remediation hours, followed by Fresno, Boston, Portland and Seattle in that order.

Table 10
Total Program Hours, by Site

Site	Mean Hours ^a	Mean Hours as a Percentage of Maximum Available Hours, Treatments	Mean Hours as a Percentage of Maximum Available Hours, Controls
<u>Boston</u>			
Treatment	165	76% (217) ^b	
Control	182		74% (246)
<u>Fresno</u>			
Treatment	216	82% (263.5)	
Control	209		72% (291)
<u>Portland</u>			
Treatment	163	72% (227)	
Control	164		72% (227)
<u>San Diego</u>			
Treatment	187	84% (223.5)	
Control	177		74% (238.5)
<u>Seattle</u>			
Treatment	167	71% (237)	
Control	182		75% (243)
<u>All Sites</u>			
Treatment	180	77% (233.6)	
Control	182		73% (249.1)

^aExcludes randomized youths with no program participation.

^bThe number in parentheses gives the maximum available hours for the group indicated.

Table 11
Remediation Hours, by Site
for Treatments Only

Site	Mean Hours ^a	Mean Hours as a Percentage of Maximum Available Hours of Remediation
Boston	93	86% (108) ^b
Fresno	96	87% (112)
Portland	87	73% (119)
San Diego	110	93% (117.5)
Seattle	83	69% (120)
All sites	94	82% (115.3)

^aExcludes randomized youth with no program participation.

^bThe number in parentheses gives the maximum available hours for the group indicated.

When the number of classroom hours was regressed on background characteristics, several variables emerged as significant predictors (Table 12). One of the more powerful predictors of classroom hours attended was a participant's baseline report of the number of times he or she was absent from school in the preceding year. A participant who had good attendance during the regular school year was likely to have good attendance in STEP's classroom component as well; youth who had poor attendance during the regular school year also had poor attendance during the summer. Participants, then, behaved toward STEP's remediation component in much the same manner as they had toward school. Being compensated at minimum wage, and having receipt of that compensation contingent upon attendance in remediation classes, was not sufficient to alter the typical attendance patterns of these youngsters.

This fact notwithstanding, it should be noted that across all categories of youth, there was a very high rate of program attendance. On average, participants attended 82% of all available classroom hours. Thus, while there may have been a relationship between absences in school and absences in remediation sessions, the rate of attendance nevertheless seems higher than might be expected in a program requiring remediation.

With school year attendance controlled, a number of additional variables influenced the amount of remediation a youth received. Those with frequent attendance at remediation sessions tended to be in the higher grades for this age range and to come from one of the minority groups -- blacks, Hispanics or Asians -- represented in the sample.

ACADEMIC PERFORMANCE

The following sections summarize findings for a number of measures of academic performance. These are findings that have emerged after a period of not more than eight weeks -- only the first phase in a 15-month treatment. Because this is a longitudinal study -- treatment and control group youth will be followed for a period of five years after completion of the STEP program -- there will be a number of additional opportunities to examine the longer-term benefits of participation in STEP. Indeed, STEP's ultimate aims are long-term ones: to reduce adolescent parenting, increase high school retention and graduation, and enhance early labor force experiences.

General Findings

In order to make a determination of program impact -- controlling for pre-treatment differences in the characteristics of the treatment and control groups -- a multiple regression model was employed. The equation controlled for baseline differences in

Table 12

Determinants of Remediation Hours

Explanatory Variable	Coefficient	Explanatory Variable	Coefficient
Constant	39.26 (t=2.20)**	Repeated a Grade	0.79 (t=0.36)
Sex (male = 1)	-0.90 (t=-0.51)	Current Grade	3.22 (t=2.35)**
Race (Black = 1)	9.53 (t=2.90)**	School Absences	-6.74 (t=-3.21)**
Race (Hispanic = 1)	7.70 (t=1.95)*	Desire for College	0.59 (t=0.29)
Race (Asian = 1)	15.92 (t=3.54)**	Baseline Reading	0.02 (t=1.20)
Public Assistance Receipt	-0.90 (t=-0.48)	Boston ¹	13.07 (t=4.37)**
Family Size	0.61 (t=1.32)	Fresno ¹	15.51 (t=5.13)**
Father Absent from Household	1.63 (t=0.85)	Portland ¹	4.18 (t=1.48)
Birth Order	-0.25 (t=-0.70)	San Diego ¹	24.85 (t=8.72)**
Limited English Proficiency	0.35 (t=0.09)		

¹The site impact is relative to the excluded category, Seattle.

* Significant at the .10 level.

** Significant at the .05 level.

R-Squared = .21

F-Test of Overall Significance = 8.39**

age, race, sex, site, English proficiency, family size, birth order, public assistance receipt, academic performance and a number of variables related to previous school attendance and performance. The dependent measures are MAT scaled scores -- the result of the conversion of the number of correct answers to a metric more suited to statistical analysis -- in reading and math.

Throughout the analysis of academic performance two methods of presentation of the data will be utilized, the choice between them to be dictated by the goals of the discussion. The first is the presentation of simple unadjusted means that depict actual scoring without statistical adjustments for any existing pre-treatment differences between treatment and control group youth on important variables. Simple means are descriptive measures that give the reader a concrete feel for the data, but are not in and of themselves definitive.

The second method -- presentation of coefficients from the regression analysis -- is more definitive precisely because statistical adjustments have been made. In the learning analysis the coefficient is a net impact score -- the difference between treatment and control youth on MAT scores at post-test, controlling for differences at baseline. Thus, a net impact score of +10 in math would mean that at post-test, treatment youth scored 10 points higher in math than their control group counterparts if their pre-treatment characteristics were identical. For reference, 20 points is roughly equivalent to a gain of half a grade in math for youth in the achievement range served by STEP; 10 points represents a gain of half a grade in reading.

Employing the multiple regression model with data from the 1,144 youth in the aggregate sample for whom both pre-test and post-test MAT scores were available and for whom data were available for all the variables in the learning regression, produced the following results with respect to treatment and control group differences:

Table 13
Net Impact of Participation
in STEP on MAT Scaled Scores in Reading and Math

	READING		MATH	
	<u>Coefficient</u>	<u>t-ratio</u>	<u>Coefficient</u>	<u>t-ratio</u>
Aggregate Sample	5.68	2.00**	12.64	3.33**

** Significant at the .05 level

Thus, in the first six to eight weeks of the program, participation in STEP had a significant impact on the academic performance of target youth. Across all sites, treatment youth scored significantly higher at post-test, net of baseline differences, in both reading and math than did control group youth. In reading, the net impact was an increase of 5.68 points; in math, treatment youth scored 12.64 points higher. In both cases, and with most of the learning effects that are to be described below, these net effects indicate that STEP is helping to stem learning losses -- e.g., reducing the magnitude of the losses that would have occurred in the absence of the treatment.

Consistent with Barbara Heyns' demonstration of summer learning losses among disadvantaged youth, STEP control group youth experienced significant losses in their academic performance over the course of the summer. As will be seen later, the aggregate sample of treatment youth also experienced losses. However, because of their participation in STEP, their losses were smaller than those of the control group. There were also subsets among treatment youth that experienced actual gains. For the aggregate sample, these subsets included Asians whose pre- to post-test scores displayed gains in both reading and math. In San Diego, gains in reading and math were found in three of the five remediation sites -- Crawford, Kearny and San Diego high schools.

Since in most instances there were no gains for treatment youth from pre-test to post-test, the significant impact of participation in STEP was a matter of decreasing the magnitude of losses that would have occurred in the absence of the treatment. Thus the validity of the finding of positive impacts depends critically upon the existence of substantial losses by control group youth. In the sections that follow, both the failure to achieve treatment gains and the validity of the control group losses are explored.

The data in Table 14, which present pre-test and post-test scores in reading and math on the Metropolitan Achievement Test (MAT) for both the aggregate sample and the individual sites, illustrate this situation in greater detail. In contrast to the data presented in the above discussion, this table presents unadjusted means. Focusing first on reading scores in the aggregate sample, one finds a significant loss of 14 points from pre-test to post-test for control group youth. Treatment youth also experienced a loss, but it was of a lesser magnitude than that of the control group, and it was not statistically significant.

When these pre-test and post-test changes are compared, one finds that participation in STEP has significantly stemmed learning losses. That is, the treatment loss of 8.55 points was significantly smaller than the control group loss of 14.28 points. In math, the pattern is similar -- participation in STEP has stemmed losses that would have occurred in the absence of the program. Control group youth experienced a significant loss of 16.47

Table 14

Pre-test to Post-test Changes in
Reading and Math Scores, by Group

Site	READING SCORES			MATH SCORES		
	Pre-Test	Post-Test	Change ^a	Pre-Test	Post-Test	Change ^a
<u>Full Sample</u>						
Treatment	717.31	708.86	-8.55	690.21	689.97	0.04
Control	718.95	704.62	-14.32 (t=2.07)**	695.96	679.16	-16.47 (t=4.40)**
<u>Boston</u>						
Treatment	733.62	724.46	-9.16	688.64	691.67	3.03
Control	722.43	694.64	-27.79 (t=2.89)*	690.25	648.89	-41.36 (t=5.81)**
<u>Fresno</u>						
Treatment	706.49	704.07	-2.87	690.73	686.41	-4.32
Control	720.21	708.88	-11.33 (t=-1.77)*	695.26	669.73	-25.53 (t=2.84)**
<u>Portland</u>						
Treatment	718.28	711.00	-7.28	683.80	690.70	8.39
Control	726.43	723.26	-3.17	696.68	697.47	0.79
<u>San Diego</u>						
Treatment	727.57	712.88	-14.69	724.84	714.73	-10.11
Control	726.95	715.67	-11.28	732.09	721.60	-10.49
<u>Seattle</u>						
Treatment	699.83	691.14	-8.69	662.89	666.34	3.45
Control	699.56	683.61	-15.95 (t=1.31)*	666.09	660.65	-3.93

^a Asterisks in this column indicate whether the change in pre-test and post-test scores is significantly different for treatments and controls.

* Significant at the .10 level, one-tailed test.

** Significant at the .05 level, one-tailed test.

points over the course of the summer, while treatment youth showed virtually no change from pre-test to post-test.

Data for the individual sites sustain the general pattern. The treatment group experienced losses of a lesser magnitude than those of the control group in Boston and Fresno in both reading and math, and in Seattle in reading. As Table 14 shows, however, there was variation among sites as to whether (a) the control group losses were significant and (b) the difference in the changes experienced by the treatment and control groups was significant.

In Boston, participation in STEP significantly stemmed learning losses in both reading and math. In Fresno, the difference in treatment and control group changes was significant at the .05 level in math, and the .10 level in reading. In Seattle, there was a significant ($p < .10$) reduction of learning losses in reading but not in math.

In Portland and San Diego, different patterns emerge. In both of these sites the treatment loss in reading was somewhat larger than that of the control group, but not significantly so. In math, San Diego treatment and control group youth experienced virtually the same degree of loss, while Portland treatment youth registered a gain against virtually no change for the control group. Like the reading result, the treatment/control group differential in math was not significantly different in either site.

The fact that participation in STEP did not result in actual gains for treatment youth is very disappointing. Since treatment youth participating in the 1984 pilot of the STEP program experienced moderate gains over the course of the summer (Branch, et al., 1984), the failure to find similar gains in 1985 was unanticipated. However, as Chapter III makes clear, STEP's remediation component was implemented with great difficulty in 1985. To recapitulate: there were problems with both the content and the timing of the development of CORE. In spite of prodigious efforts on the part of both program planners and operators, at a number of sites CORE failed to project a sense of being "something special," different from school. Moreover, the remediation component was not adequately linked to the life skills or work experience components of the program. With respect to logistics, the curriculum development process began late -- thus remediation teachers were not identified or trained early enough to become sufficiently familiar with the self-paced, competency-based format of CORE and appropriate materials were often not available on a timely basis or in the required quantities. Thus, there is sufficient explanation within the implementation of the remediation component itself to explain the lack of treatment gains.

The Case of San Diego

It is instructive, however, to explore the experiences of San Diego treatment youth to better understand the failure to achieve gains. In San Diego, participants attended remediation sessions at five different high schools -- Crawford, Kearny, Lincoln, Morse and San Diego. Table 15 presents changes in San Diego scores by remediation site. In three sites, there were gains for the treatment group in both reading and math -- fairly sizable gains in reading at San Diego and Kearny -- and in math at Crawford and Kearny. At Lincoln and Morse high schools, treatment youth experienced dramatic losses -- the largest for any group of treatments or controls in the demonstration. As the discussion that follows implies, the potential of the model to produce gains was demonstrated by the three San Diego sites that operated under more optimal conditions.

Since the San Diego remediation sites produced both the largest gains and the largest losses for treatment youth, it would be useful to understand what accounts for the differences among them. First, ethnic distribution varied among the sites. Participants were assigned to remediation sites primarily on the basis of residence. Thus Lincoln and Morse (the schools with the poorest outcomes for both reading and math) were primarily black -- 86 percent and 79 percent respectively, while Kearny (the school with the best reading and math outcomes) was primarily Asian (71%). San Diego was racially mixed (30% black, 42% Hispanic, 23% Asian) and Crawford had roughly equivalent numbers of blacks and Asians.

In addition to the racial differences, differences in the physical environments of the sites and in the availability of computers were documented by systematic observations of remediation sessions. Kearny (and San Diego, where the learning impacts were more mixed) had a clearly superior physical environment; in addition, there were twice as many computers available at Kearny as at any other site -- enough for individual access by each student. Lincoln and Morse, on the other hand, had poor physical environments -- older, rundown buildings, crowded rooms and poor lighting. The number of computers, while dramatically fewer than at Kearny, was about the same as at San Diego or Crawford.

Physical differences notwithstanding, the data do not document major differences in the instructional characteristics of the sites. In all sites, the majority of youth appeared to be attentive and engaged in the remediation and, with one exception, student and teacher attitudes toward each other seemed uniformly neutral. While it is conceivable that the quality of instruction at Lincoln and Morse was lower than that at other sites, observations of remediation sessions were not sufficient to make such a judgment. It is known, however, that perhaps the most able teacher in the program -- a math teacher who had assisted in the development of the curriculum -- taught at Kearny. Test condi-

Table 15

Pre-test to Post-test Changes in Reading and Math Scores
in San Diego, by Remediation Site

Site	READING SCORES			MATH SCORES		
	Pre-Test	Post-Test	Change ^a	Pre-Test	Post-Test	Change ^a
<u>San Diego</u>						
Crawford High	745.04	746.77	1.73	740.04	778.62	38.62
Kearny High	725.42	734.27	8.84	756.46	783.19	26.73
Lincoln High	721.09	663.16	-57.94	700.97	643.81	-57.16
Morse High	715.54	687.31	-28.23	691.15	643.81	-47.35
San Diego High	732.67	747.00	14.33	742.50	742.75	0.25
			(F=8.37)**			(F=10.14)**
All Sites	726.95	715.67	-11.28	732.09	721.60	-10.49

^aAsterisks in this column indicate whether the change in pre-test and post-test scores is significantly different among the subsites.

**Significant at the .05 level.

tions, possibly, offer additional insights. Post-program interviews suggest that some remediation teachers may have transmitted negative attitudes about the test to their students.

While the difference in learning outcomes among the San Diego remediation sites seems in part due to differences among racial groups in improvement over the summer, it should be noted that black treatment youth in San Diego had larger losses than black treatment or control youth at any other site. In addition, the quality of remediation facilities and the attitudes and ability levels of the remediation teachers appear to have contributed to the observed differences.

Control Group Losses

While the failure to find treatment gains can be explained by problems in the implementation of STEP's remediation component, understanding control group losses is a more complex issue. Heyns' original research found summer learning losses among low-income youth in Atlanta, resulting from the lack of instruction and daily practice during the summer. STEP's pilot research (Branch, et al., 1984) also found significant control group losses. In both cases, however, the possibility that these losses were an artifact of the control group's lesser motivation to perform well on the post-test could not be ruled out.

In order to minimize (if not eliminate) the likelihood that control group youth's lack of motivation was spuriously elevating the estimate of program impact, as mentioned in Chapter II, an incentive scheme was devised to provide both treatment and control group youth with an external reason for doing their best at both pre-test and post-test. Thus, at each testing session, participating youth received one chance in a drawing for one of 15 \$50 prizes at each site to be awarded for taking the test. They received an additional chance for each correct answer scored. If these incentives had the desired effects, any observed control group losses should be the result of valid summer learning losses.

Even with this incentive, moderate control group losses emerged for the aggregate sample and for most of the individual sites, lending support to the belief that they are reflective of actual losses in reading and math skills. The magnitude of the control group losses are not uniform, however; Boston's control group losses in both reading and math appear particularly large. Indeed, when a regression model is used to determine if these site differences in control group post-test scores are significant, one finds that, net of the influence of a number of variables including baseline score on the MAT, the size of Boston's control group losses in reading is significantly larger than those at Portland and Fresno, and in math is significantly larger than for any other site except Fresno.

This variation in the magnitude of control group losses is somewhat troublesome, in that it may imply irregularities in post-test administration. Scrutiny of the post-testing procedures at the sites revealed a degree of variation that, while greater than could be desired from a research perspective, is (a) consistent with what one can reasonably expect in field research and (b) does not logically seem responsible for the observed differences in the size of control group losses. For instance, treatment and control youth were post-tested separately at all the sites, not just in Boston. It was equally possible at any of the sites to introduce differential motivation to do well, but there is no reason to believe that it occurred in any of the sites.

Alternatively, one might focus on the Portland findings. Control group youth in this site experienced a small (-3.17) loss in reading and virtually no change (+0.79) in math. Why were there no control group losses? In part, this phenomenon is a function of racial differences in Portland. Asian and white control group youth experienced gains in reading and math, while black and Hispanic control group youth experienced losses. Racial differences in program impact are discussed in greater detail in the sections that follow.

Site variation notwithstanding, the emergence of significant control group losses in the face of external incentives to increase performance has strengthened the validity of the learning losses phenomenon. A number of additional investigations were undertaken to further explore this question. These analyses, which are presented in full in the Appendix, seek to determine whether the post-test data for controls are less reliable than the same data for treatment youths. Analysis of a number of measures of internal consistency and reliability -- split half measures and the Kuder-Richardson (KR20) measure of reliability -- were generally not consistent with a hypothesis of increased random answering on the post-test by controls. Since these tests, for the most part, did not point to a decline in post-test reliability for controls that was untrue of treatments, confidence that measured declines in test scores are reflective of true learning losses is increased.

Subgroup Analyses

The preceding discussions have focused on average treatment effects. In fact, as Table 16 shows, the impact of participation in STEP varied as a function of sex, race, site and other key variables.

Sex. Participation in STEP had a differential effect on boys and girls. Treatment girls scored approximately 11 points higher than control group girls on the reading post-test, while the scores of treatment and control group boys were virtually the same. In math, however, both boys and girls in the treatment

Table 16

Net Impact of Participation in STEP as
Revealed by Multivariate Analysis of MAT Scaled Scores
in Reading and Math

	READING		MATH	
	Coefficient	t-ratio	Coefficient	t-ratio
Full Sample	5.68	2.00**	12.64	3.33**
<u>Sites</u>				
Boston	22.47	3.20**	39.46	4.60**
Fresno	4.79	1.00	19.97	2.68**
Portland	-12.43	-1.63*	-5.53	-0.56
San Diego	5.42	0.81	0.13	0.02
Seattle	9.39	1.58*	5.52	0.65
<u>Sex</u>				
Female	11.37	3.18**	12.40	2.68**
Male	0.45	0.10	12.04	1.98**
<u>Race</u>				
Black	0.43	0.11	11.03	2.25**
Hispanic	17.18	2.46**	20.73	2.16**
Asian	10.19	1.67**	6.61	0.73
White & Other	11.90	1.09	14.64	0.98

NOTE: This table presents selected results from 24 separate regressions, one each for reading and math for the full sample and each of the subgroups presented in the table.

* Significant at the .10 level, one-tailed test.

** Significant at the .05 level, one-tailed test.

group significantly outscored their control group counterparts -- by about 12 points for both.

Site. The impact of participation in STEP varied among the sites. In both reading and math, treatment youth in Boston scored significantly higher than control group youth at post-test -- by 22 and 39 points respectively. In Fresno, treatment youth scored 20 points higher than control youth in math, and in Seattle treatment youth significantly ($p < .10$) outscored control group youth in reading.

It is clear, however, that treatment youth in Portland did not fare as well vis-a-vis their control group counterparts as participants at other sites. In fact, in reading, control group youth significantly ($p < .10$) outscored treatment youth. This was largely a function of the fact that controls in Portland experienced relatively small losses. In fact, Asian and white controls in Portland actually had gains. Since their scores serve as a proxy for what would have happened in the absence of the program, if STEP had a positive effect one would expect even larger gains for treatment youth. The losses experienced by treatment youth, then, are clearly a reflection of the implementation difficulties experienced at this site. Portland, unlike other sites, had the task of implementing a new program -- STEP -- while simultaneously undergoing a system-wide reorganization of the Summer Youth Employment and Training Program. This exacerbated the implementation difficulties that occurred demonstration-wide.

Race/Ethnicity. The impact of participation in STEP also varied among the racial/ethnic groups represented in the samples. Considering the net impact on reading, it is clear that Hispanics and Asians benefited most from participation in STEP. Hispanics in the treatment group scored 17 points higher on reading at post-test than Hispanics in the control group. Asian youth in the treatment group scored on average 10 points higher in reading than Asian youth in the control group. The impact of participation in STEP on the reading scores of blacks, whites and others was not statistically significant.

Hispanic youth also appear to be the group whose performance in math profited most from participation in STEP: at post-test, Hispanic youth in the treatment group scored 21 points higher than Hispanic youth who did not participate in STEP. Black treatment youth also gained in math, relative to black control group youth -- on average, by 11 points.

As Table 17 shows, there were noteworthy differences in the way learning impact was attained by Hispanic, black and Asian youth. For Hispanic youth it is clearly a matter of stemming losses: while Hispanic youth in the treatment group experienced losses in reading and math of six and eight points respectively, their control group counterparts experienced more dramatic losses of 28 and 35 points. Thus STEP's strong impact for Hispanics is a

Table 17

Pre-test to Post-test Change in Scores, by Race

	READING SCORES			MATH SCORES		
	Pre-Test	Post-Test	Change ^a	Pre-test	Post-test	Change ^a
<u>Treatments</u>						
Black	714.85	701.92	-13.10**	672.09	669.58	-2.08
Hispanic	725.54	719.35	-6.19	710.92	702.93	-7.99
Asian	713.13	721.06	7.93*	745.78	760.47	14.69**
White & Other	729.41	720.00	-9.41	684.71	688.49	3.78
All	717.31	708.86	-8.55	690.21	689.97	0.04
<u>Controls</u>						
Black	718.06	705.09	-12.97**	682.95	665.66	-16.72**
Hispanic	725.38	697.61	-27.77**	708.34	672.95	-35.39**
Asian	712.99	707.72	-5.27	736.64	736.56	-0.07
White & Other	722.04	707.43	-14.61**	691.88	680.39	-11.49
All	718.95	704.62	-14.32	695.96	679.16	-16.47
<u>Total Sample</u>						
Black	716.36	703.41	-13.04**	677.19	667.74	-8.96**
Hispanic	725.46	708.59	-16.87**	709.64	688.09	-21.55**
Asian	713.06	714.62	1.56	741.37	748.93	7.57*
White & Other	724.85	712.22	-12.63**	689.15	683.47	-5.68
All	718.11	706.78	-11.39	693.03	684.66	-8.07

^a Asterisks in this column indicate whether the change between pre-test and post-test scores is significantly different from zero.

* Significant at .10 level.

** Significant at .05 level.

function of having been able to diminish dramatic losses that would have occurred otherwise. Much the same is true for blacks. Their significant math impact is a function of a small treatment loss contrasted with a more substantial loss for the control group. For Asian youth, however, control group youth did not experience large losses. In reading, they lost only five points and in math their scores remained unchanged. The significant net impact in reading for Asian youth, then, is due to a significant treatment gain (7.93 points) over the summer.

Site and Race. As Chapter IV's sample profile suggests, it is difficult to separate the effects of site and race. Most of the sample's Hispanics come from Boston, Fresno and San Diego. Since two of the three -- Boston and Fresno -- are sites with significant treatment effects, it is important to ask which of these factors is driving the impact. Is the Hispanic effect really a Boston effect, or do the Hispanics in Boston (primarily Puerto Rican) and the Hispanics in Fresno or San Diego (primarily Mexican-American) fare equally well?

In order to address this question, data from the five sites were analyzed, using a multiple regression model that includes terms interacting treatment with site and race. These terms assess the possibility that the impact of participation in STEP will vary as a function of a participant's site or race. The principal findings from this analysis appear in Table 18, which presents the coefficients from those sex and race combinations that for each of the sites was found to be significant.

What these data illustrate with respect to the site/race question is, first, that in Boston the net impact of participation in STEP was significant in both reading and math for all races, not just Hispanics. Second, in Fresno, San Diego and Seattle there was a significant treatment effect for both reading and math for Hispanic youth. This effect applies to Hispanic females in all three sites, and to Hispanic males ($p < .10$) as well in Seattle. In math, there was a significant impact of participation on the scores of Hispanic males and females in Fresno, but none of the other sites. Thus it appears that in math, at least, there were separate effects for race and site.

In reading, there was clearly a significant impact for Boston; the Hispanic effect depended upon the sex of the participant -- Hispanic females significantly outscored their control group counterparts in Boston, Fresno and San Diego. However, it is also clear that black youth benefited least from participation in STEP, and of them, black males had particularly distressing outcomes. In two sites -- Portland and San Diego -- black males fared significantly worse in reading than their control group counterparts.

An examination of the means in Table 17 reveals a number of additional findings worth discussing. In contrast to Heyn's study,

Table 18

Statistically Significant Impacts in Pooled
Regression Model with Multiple Interaction Terms

Site	Race	Sex	Reading Impact	Math Impact
Boston	Black	Male	13.1*	42.3
		Female	24.0	42.4
	Hispanic	Male	32.7	48.4
		Female	43.7	48.5
	Asian	Male	29.1	54.7*
		Female	40.1	54.8
	White & Other	Male	22.2	37.3
		Female	33.1	37.4
Fresno	Hispanic	Male		24.5
		Female	20.1	24.6
	Asian	Male		30.8
		Female		30.9*
Portland	Black	Male	-21.4	
San Diego	Black	Male	-13.1	
	Hispanic	Female	17.5*	
Seattle	Hispanic	Male	19.3*	
		Female	30.3	
	Asian	Male	15.7*	
		Female	26.7	
	White & Other	Female	19.8*	

*Significant at the .10 level. All other table entries are significant at the .05 level.

NOTE: Entries in this table represent impact of treatment on type of individual specified. This is computed as the sum of base treatment coefficient plus each appropriate treatment interaction term coefficient. Significance is then determined by appropriate F-test of the hypothesis that this sum equals zero.

large differences did not emerge in the performance of black and white youth. White youth in the control group experienced losses in both reading and math of similar magnitude to those of black control group youth. Among treatment youth, blacks and whites experienced similar losses in reading. In math, however, there was a small difference -- blacks experienced a non-significant loss of two points while whites experienced a non-significant gain.

The most striking racial difference in the sample involves the academic performance of Asian youth. Their control group losses were smaller than those of any other youth; in fact, in math, there was virtually no change (-0.07) from pre-test to post-test for them. Among treatment youth, Asians experienced significant gains in both reading (7.93) and math (14.69). It does not appear that known differences among the races in certain indices of socioeconomic status are useful in understanding these differences. All youth meet JTPA eligibility standards that certify low income status, and while there are differences among the races in the percent coming from two-parent households, this variable cannot explain the differences in academic performance.

Baseline Ability Groups. It is also of interest to ascertain whether STEP was able to produce gains for all participants regardless of how well they scored on baseline measures of academic performance or whether program impact was greater for either the most able or the least able among them. Since STEP's remediation component is individualized and self-paced -- and students begin at a point consistent with their abilities and proceed at their own pace -- one might hypothesize that there would be less variation than in other instructional modes.

Alternatively, one might hypothesize that an individualized mode of instruction might make it easier to focus on one ability group versus another. In order to investigate this possibility, the sample was divided into quartiles (four groups of equal numbers of youth) according to baseline MAT scores, and an analysis was undertaken to determine if relative differences in program impact existed among them -- if STEP had produced improvements for the top quartile that were substantially different from those produced for the bottom quartile.

Table 19 presents the data on the pre-test to post-test change scores for youth in the various quartiles. The interpretation of these data is not straightforward. A potential source of difficulty is the phenomenon of regression toward the mean. Because no test is perfectly reliable -- how one scores can be influenced by luck, environmental conditions or simply having an off day -- people who score extremely low on a test's first administration will tend to score closer to the mean on subsequent administrations. The same applies for people scoring extremely high on the first administration; without any change in the trait a test measures, there will be a tendency to shift toward the mean on the

Table 19

Program Impact According to
Pre-Program Achievement Levels

Group & Quartile*	Lowest Quartile	Second Quartile	Third Quartile	Highest Quartile
Treatment Reading	8.1	-13.4	-9.2	-20.1
Control Reading	-0.8	-17.1	-13.0	-27.4
Difference	8.9	3.7	3.8	7.3
Treatment Math	19.4	4.6	-3.1	-24.0
Control Math	-5.2	-16.2	-15.2	-31.6
Difference	24.6	20.8	12.1	7.6

*The quartiles are the result of ranking all STEP youth according to pre-test reading score and then dividing them into four equal groups.

NOTE: Table entries represent the change in MAT test score from pre-test to post-test for the group indicated.

second administration. This makes it difficult to isolate a differential impact of treatment among the various ability levels.

For example, one could argue that in Table 19, the math results for treatment youth indicate that STEP was more effective for the least able among them, since youth in the lower quartiles experienced the largest gains while those in the upper quartiles experienced the largest losses. An alternative hypothesis is that this trend is the result of regression toward the mean. If this were true, one would expect the largest gain to be experienced by youth in the lowest quartile and the smallest by youth in the highest quartile. An examination of the control group results tends to support this hypothesis. Control group youth showed the same trend -- smaller losses for the least able youth and larger losses for the most able -- without benefit of treatment.

As a final approach to the investigation of this question, the multiple regression model for the math equation was modified to include a number of terms that allow for the interaction of treatment and baseline ability levels as measured in quartiles. An F-test was then performed to test the null hypothesis that the interaction terms were simultaneously zero. The results of this analysis appear in Table 20. Since this hypothesis could not be rejected, the evidence does not suggest that STEP had a differential impact on the participants based on their ability levels. A similar analysis was undertaken for reading, with similar results.

Hours Analysis

Thus far, the impact of STEP has been discussed as an all-or-nothing proposition -- eligible youth either received the STEP treatment or not. The estimate of the treatment impact derived from such an analysis is not sensitive to differences in the amount of remediation actually received. Included in this estimate of program impact are data both from treatment youth who dropped out of the program after relatively little exposure to remediation and those who attended the program conscientiously and received most of the planned remediation. However, since it is reasonable to hypothesize that youth who are exposed to more remediation are likely to have better outcomes, an investigation of the incremental impact on learning of additional hours of remediation was undertaken.

In order to correct for the possibility that some of the same variables (for instance, race) that predict amount of remediation received also predict how a youth is likely to do on the post-test, a system of three equations (hours, reading and math) were estimated using the two-stage least squares technique.

Table 20

Math Impact with Treatment Interacted with
Ability Levels in a Pooled Regression

Explanatory Variable	Estimated Coefficient	Explanatory Variable	Estimated Coefficient
Constant	19.57 (t=0.34)	Reading pre-test	0.27 (t=3.44)
Male	-16.13 (t=-4.20)	Math pre-test	0.58 (t=16.31)
Black	-10.40 (t=-1.60)	Enjoys math (base)	-10.65 ³ (t=-5.98)
Hispanic	-9.77 (t=-1.20)	Treatment	18.01 (t=2.37)
Asian	27.28 (t=2.88)	Boston ¹	1.92 (t=0.30)
Receives AFDC and/ or Food Stamps	0.93 (t=0.23)	Fresno ¹	4.54 (t=0.67)
Family Size	-0.78 (t=-0.76)	Portland ¹	11.00 (t=1.78)
Lives Without Father	3.28 (t=0.78)	San Diego ¹	10.72 (t=1.67)
Position in Birth Order	0.35 (t=0.44)	Treatment/2nd Quartile Interaction Term ²	4.85 (t=0.45)
Problems with English	6.43 (t=0.79)	Treatment/3rd Quartile Interaction Term ²	-6.87 (t=-0.64)
Repeated a Grade	3.70 (t=0.77)	Treatment/4th Quartile Interaction Term ²	-19.09 (t=-1.79)
Grade last year	10.81 (t=3.63)	Scored in Second Quartile	-4.21 (t=-0.50)
Absence Often in Past School Year	4.57 (t=1.02)	Scored in Third Quartile	7.07 (t=0.72)
Expects Post-High School Education (base)	3.23 (t=0.72)	Scored in Fourth Quartile	1.50 (t=0.11)

R-squared = 0.53

F-test 46.18

¹The site impact is relative to the excluded category -- Seattle.²F-test = 1.88 for the hypothesis that the treatment interaction terms were simultaneously zero.³This variable was scaled so that the lowest score was associated with the highest level of enjoyment.

For the aggregate sample, the following results were generated:

	<u>Reading Score</u>	<u>Math Score</u>
Impact per Treatment Hour	1.1**	-0.6

**Significant at the .05 level.

As they indicate, all things being equal, an additional hour of classroom participation would result in an increase of 1.1 points in a treatment youth's post-test score in reading and no significant impact in math.

This is a fairly sizable impact when one remembers that the hours variable is not a measure of "time-on-task." Rather, it is a much less specific measure that includes life planning hours as well as remediation hours, both reading and math instruction, breaks and distractions, as well as time specifically devoted to remediation. Under these circumstances, a one point increase per hour of classroom time is a substantial one. If it were possible to increase a youth's classroom time by 10 hours (roughly an additional three days of remediation) -- by developing effective incentives, for example -- one might expect a youth to score an average of 10 points higher on the MAT.

In the preceding analyses, the incremental impact of additional learning hours of remediation was examined while controlling for the jointly determined nature of the hours. One might ask a different question: How does the program work for treatment youth who are motivated enough to attend frequently and complete the program? Do these youth attain actual gains from pre-test to post-test? When the regression model is estimated for the sample of 1,073 youths who did not drop out of the program at any point, one finds that treatment youth outscore control group youth by 8.14 points in reading and 15.80 points in math. When one uses these regression results to predict post-test scores for this sample, one finds a small non-significant gain (from 695 to 697) in math and a loss (from 720 to 715) in reading. Thus, while the program had a stronger impact on those youth who complete it, it was not strong enough to result in an actual gain.

RETURN TO SCHOOL

As a matter of policy, students under the age of 16 cannot officially drop out of school. Therefore, it is not surprising that the overwhelming majority of the youth in the STEP sample -- both treatments and controls -- returned to school at the end of the summer.

Of the youth for whom data were available, 92 percent had returned to school. The 8 percent of youth who were out of school in-

cluded those who had taken full-time jobs, were in detention or were no longer in contact with their families. These figures did not vary by treatment or control group status -- 92 percent of treatments and 91 percent of controls were in school again in the fall. Nor were there site, race or sex differences of note. At this point in their educational careers, the return to school rate for these young people is fairly high. It is over the next several years -- particularly as the sample reaches the age of 16 -- that increasing numbers of youth will begin to drop out of school. Differences between treatments and controls are likely to emerge at this point.

LIFE SKILLS OUTCOMES

In addition to hypotheses concerning the impact of STEP on academic performance, it was also predicted that participation in STEP would lead to a number of benefits associated with the life skills component. These benefits are related to the general theme of the curriculum: in order for participants to prepare for their futures, they must begin now to make critical decisions about education, work and how to respond to pressures for sexual involvement. In the sections that follow, program impact on educational aspirations, career orientation, and fertility-related attitudes and behavior are explored.

Educational Attitudes

Members of the sample were asked a number of questions measuring their educational aspirations, expectations and perceived abilities. Specifically, they were asked how much they enjoyed reading and math; whether they considered themselves better, worse or about the same as their peers in reading and math; how important a high school education was to them; and whether and how far they intended to pursue education after high school. Changes in such attitudes, it has been hypothesized, would pave the way for improvement in academic performance and a greater attachment to school.

Very little change, however, was found on any of these measures. For the measures of educational aspiration, this seems due to lack of variance. Ninety-nine percent of treatment youth and 98% of controls said, at baseline, that graduating from high school was important to them. When the follow-up questionnaire was administered six weeks later, there had been no shift in these attitudes. Similarly, none of these young people (less than one percent) thought it likely that they would leave high school without receiving a diploma; indeed only 22 percent of them believed that they would not go on to college. Again, there was no major shift over the course of the summer in these percentages.

Nor were there shifts in participants' enjoyment of reading and math, or their evaluations of their abilities to perform well in these subject areas. Although there was no change over the

course of the summer, the percent saying that they enjoyed math (51%) was noticeably higher than those who said they enjoyed reading (37%). As for perceived academic abilities, for both reading and math, 25 percent of the sample felt they were better than their peers. These percentages remained stable throughout the summer.

The examination of these data has not revealed significant changes in any of the educational attitudes of sample youth over the course of the program. This is due at least in part to the shortness of the observational period and to the high expectations that these young people hold for their own academic futures. This suggests that STEP is correctly targetted; it has reached a group of disadvantaged youth at risk of dropping out of school, before they have lost their optimism about their educational prospects. Their expectations stand in sharp contrast to what the research literature predicts. A better test of whether STEP has a positive impact on these attitudes can be had over the long term. It could be that treatment youth will be better able to maintain these expectations and indeed convert them into concrete behavior, as their control group counterparts steadily decline, both in expectations and actual school enrollment.

Career Orientations

One aspect of the life skills curriculum is devoted to expanding participants' awareness of their future work lives, and the steps that must be taken to prepare for careers. One of its segments focused on career options open to women, with the goal of influencing young women to pursue non-traditional careers. It was hypothesized that STEP participants would be more likely than members of the control group to consider it appropriate for women to pursue a wide variety of traditional and non-traditional careers.

One series of questions asked youth to indicate whether they thought it appropriate for women to hold each of a total of 18 jobs -- among them those of truck driver, doctor, secretary, teacher and bookkeeper. Counter to the hypothesis, STEP participants did not consider more jobs appropriate for women than did members of the control group. Nor were there extensive differences by age or race. When separate regression analyses were conducted, a significant impact was found for boys. Treatment boys at the end of the summer believed that more jobs were appropriate for women than did control boys. The number of jobs that girls -- both treatments and controls -- thought appropriate for women was so uniformly high that there was too little variation to be explained in a multivariate analysis.

FERTILITY-RELATED OUTCOMES

For many disadvantaged youth -- particularly young women -- adolescent pregnancy or parenthood is a major obstacle to high

school completion. It was hypothesized that STEP participants, because of their exposure to both academic remediation and life skills instruction, would be less likely to drop out of school over the long run because of adolescent pregnancy and parenthood. Because of the shortness of the summer, however, it was not possible to observe the impact of the program on these long-term outcomes. In their stead, variables such as knowledge of methods of birth control, awareness of the consequences of adolescent parenthood, the use of contraception among sexually active youth, and delay in the initiation of sexual activity -- intermediate variables that serve as proxies for the long-term outcomes -- have been analyzed.

The sections that follow present the results of these analyses. The analysis of birth control knowledge and awareness of the consequences of adolescent parenthood is presented first.

Birth Control Knowledge

STEP treatment youth received 18 hours of instruction in life skills, aimed at increasing responsible social and sexual behavior. An additional part of this component was an informational tour of a comprehensive health facility -- to introduce youth to the medical services provided there and to acquaint them with a source of family planning information and assistance.

To determine the impact of this treatment on birth control knowledge, participants were asked at pre-test and post-test to respond to a series of 12 true or false statements designed to tap knowledge of the use and effectiveness of various birth control methods. As Table 21 shows, treatment youth at post-test were able to correctly answer, on average, 2.55 more questions than their control group counterparts on this measure of birth control knowledge. This is a sizable impact, representing an increase of approximately one-third in this measure of birth control knowledge.

A similar finding was obtained for a second measure of fertility-related knowledge -- in this case a series of questions testing participants' knowledge of the availability of contraception.

Youth were asked whether each of eight contraceptives was available over the counter or required a doctor's prescription. At post-test, treatment youth could correctly answer, on average, 1.73 more questions about the availability of these forms of contraceptives than could control group youth. This again represents an increase in knowledge of roughly a third.

This finding -- that participation in STEP has a substantial impact upon participants' knowledge of fertility-related information -- is general throughout the sample. As Table 22 shows, across every site, race and sex subgroup, STEP had a significant impact on participants' birth control knowledge.

Table 21

Net Impact of Participation in STEP on Birth Control
Knowledge Variables, by Site and Selected Subgroups

	Contraceptive Knowledge; Mean Number Correct		Knowledge of the Avail- ability of Contraception; Mean Number Correct	
	Coefficient	t-ratio	Coefficient	t-ratio
Full Sample	2.55	18.60**	1.73	16.42**
<u>SITES</u>				
Boston	1.71	5.21**	.88	3.49**
Fresno	3.90	12.50**	2.62	10.51**
Portland	2.19	7.30**	1.38	5.72**
San Diego	3.06	10.41**	2.36	10.97**
Seattle	1.73	4.87**	1.29	5.13**
<u>SEX</u>				
Male	2.39	11.15**	1.67	10.48**
Female	2.61	14.40**	1.75	12.15**
<u>RACE</u>				
Asian	2.58	6.32**	2.14	7.88**
Black	2.29	13.69**	1.56	11.35**
Hispanic	3.49	9.20**	2.38	9.06**
White & Other	2.85	5.30**	1.51	3.97**

**Significant at the .05 level.

Table 22
Percent of Respondents Sexually Active
In the Previous Two Months¹

	Baseline	Follow-Up
<u>Full Sample</u>		
Treatments	24.94	30.43
Controls	27.31	33.63
<u>Males</u>		
Treatments	37.62	42.08
Controls	38.94	50.88
<u>Females</u>		
Treatments	14.04	20.43
Controls	15.21	15.67

¹Based on total sample of 880 respondents for whom data on recent sexual activity was available at both baseline and follow-up.

Views of Adolescent Parenthood. The second measure of fertility-related attitudes is knowledge of the consequences of adolescent parenthood -- a two-item scale asking participants the cost of raising a child to age 18 and the likelihood that a girl experiencing a pregnancy in adolescence will finish high school.

An index measuring knowledge of the consequences of adolescent pregnancy was created from responses to these questions. Scores on this measure were then analyzed for the impact of participation in STEP, holding constant baseline knowledge of the consequences of adolescent parenthood as well as the same list of variables used for the birth control knowledge analysis. This analysis revealed that there was not a significant impact of participation in STEP on awareness of the consequences of adolescent parenthood. Treatment and control group youth experienced similar increases in their awareness of this information over the course of the summer.

Sexual Behavior. In addition to these measures of knowledge, the participant questionnaire also included a number of questions about sexual activity and the use of contraceptives. At both baseline and follow-up, participants were asked if they were sexually experienced, whether they had engaged in sexual intercourse in the past two months, whether they had used contraceptives and whether they had ever been pregnant or fathered a child.

Because upwards of 50 percent of the sample was not sexually experienced at baseline, the numbers available for analysis of key behavioral questions are relatively small. With such sample sizes, differences of even a sizable magnitude may fail to reach statistical significance. We will be in a much better position later -- participants will be followed for upwards of five years after program completion -- to examine STEP's impact on this behavior.

However, one very important change did emerge -- even in the short term -- in the sexual behavior of the young men in the sample. Participation in STEP had the effect of decreasing the likelihood that boys would engage in sex over the course of the summer. That is, at the end of the summer, treatment boys were significantly more likely than boys in the control group to say that they had not had sex in the past two months. As Table 22 shows, the percent of boys saying they had had sex during the summer did not decline. In fact, for both treatment and control group boys, there was an increase from baseline to follow-up. However, the increase was smaller for treatment boys (5%) than for control group boys (12%). Multivariate analysis of these data -- in this case a probit analysis -- confirms that the treatment/control group differences are statistically significant.

Because so few studies have documented a positive impact on the sexual behavior of youth of either gender, this finding seems surprising. Even fewer studies have documented programs that have positively affected the sexual behavior of boys. While a number of programs have changed sexual attitudes or have increased knowledge of effective birth control techniques, it is widely acknowledged that sex and fertility-control behavior is notoriously difficult to affect. Thus, it is important to scrutinize this finding both to determine that it is not an artifact, and to understand how it was achieved.

It is possible that boys in a program that counsels responsible sexual behavior would be less likely than control group boys to admit sexual activity regardless of their actual behavior. These findings, then, could be an artifact. For treatment boys, but not control group boys, denying sexual activity might be the socially desirable response.

This is an issue that can never be completely resolved, since relying upon self-report is the only practical way to study sexual behavior in a program like STEP. However, a number of observations suggest that these self-reports are credible. One of the strongest comes from examining the responses of treatment girls. They too participated in the life skills sessions but were no more likely than control group girls to give the socially desirable response -- that they had not been sexually active during the program. If social desirability was not a significant factor in the responses of treatment girls, it should not be a factor in the responses of boys. Moreover, one might think that it would not be quite so easy, no matter how strong the program norms, to get boys to admit that they were not having sex. Thus, it seems reasonable to assume that the differences in the reports of treatment and control group boys reflect actual differences in their behavior, and that it is appropriate to discuss how STEP has achieved these results.

First, STEP is one of few programs addressing the topic of pregnancy prevention in a way that includes a balanced intervention for both boys and girls. The life skills curriculum draws out the implications of irresponsible sexual behavior for both sexes. Boys, perhaps for the first time, learn that they too can be adversely affected if they father a child while still in school. They also hear about the difficulties that can result from indiscriminate sexual behavior -- difficulties in relationships and increased risk of sexually transmitted disease. These messages are conveyed through films, exercises and role play designed specifically to speak to young men.

It is not clear why there was no reduction in the number of girls who were sexually active during the summer. It is worth noting, however, that girls were considerably less likely than boys to be sexually experienced or to have been active in the prior two months. It could be that girls hear the information in the life

skills curriculum more often than boys and it therefore has less impact, particularly on the girls who have already made the decision to become sexually active at this early age.

The result in Fresno was exactly opposite. That is, treatment youth in Fresno were significantly more likely than their control group counterparts to have engaged in sexual behavior during the program period. When one examines the means, it becomes clear that it is the treatment males in Fresno who, unlike treatment males in the aggregate, are responsible for the increased sexual activity at this site.

The reasons for this anomalous result are not clear. The data do not support the hypothesis that treatment youth in Fresno are better contraceptors and, therefore, rightfully feel more confident that they can increase their sexual activity without increasing the risk of pregnancy. There is more support for the notion that the life skills component was not as successfully implemented in Fresno as in other sites. In Fresno, more than in other sites, discomfort with the life skills component was expressed by some of the teachers. Believing that the group discussion format was not appropriate for this target population, teachers in Fresno often substituted lectures and other approaches for the exercises prescribed by the curriculum.

VI. SUMMARY AND CONCLUSIONS

This chapter summarizes the findings from the first summer of the STEP demonstration, examining both its operational feasibility and its short-term impact on the academic performance and life skills attitudes and behavior of its participants.

The experience to date has been encouraging. The demonstration was implemented in accordance with the basic program design and STEP's target population was served as planned. Similarly, STEP has been successful in attaining a number of benefits for its participants. Participants fared significantly better than their control group counterparts in reading and math. In addition, participation in STEP substantially increased participants' knowledge of birth control information, and reduced the likelihood that males would be sexually active during the program period.

PROGRAM IMPLEMENTATION

General

STEP implementation requires close collaboration among a variety of institutions at a level of intensity that was new to most sites. Overall, the STEP sites were successful in achieving the collaboration essential to identifying and recruiting special target youth, securing parental consent to examine school records, hiring and training teachers, identifying appropriate curricula and materials for remediation, and meeting the tight timelines of a short summer program. Employment and training agencies, school systems and community organizations collaborated in the implementation of STEP to an unusual degree.

Local administrative arrangements varied widely, reflecting both the flexibility of the STEP model and traditional roles and responsibilities within sites. Each site allocated responsibility for planning, managing and operating STEP to a different configuration of agencies. In all sites, however, the level of commitment to STEP was high. In every case, at least one of the collaborating agencies had already been looking for ways to enrich the summer jobs program before deciding to participate in the STEP demonstration.

The public school systems were involved more than is generally characteristic of summer youth programs. As a by-product of STEP, therefore, collaboration between the schools and JTPA programming generally has been increased. The school systems in four sites "certified" the instructional component of STEP by granting credit for successful completion.

Participants

The demonstration recruited a large number of youth, clearly in need of a complex, multiple-treatment program such as that offered by STEP. By all accounts, this group of young people is highly at risk for dropping out of high school. Specifically, at 14 and 15, many have already suffered significant academic setbacks. In addition, although the majority are sexually experienced, they are not very knowledgeable about contraception, nor do they practice it with any frequency.

In spite of these objective conditions, the youths themselves are very optimistic about their educational prospects. The overwhelming majority believe that they will not only graduate from high school, but will go on to some post-secondary school experience. STEP, then, has been successful in intervening with these young people before they have become permanently alienated from school. It is important for STEP and for these young people's subsequent educators to nurture and sustain this optimism.

Participation in STEP was fairly high for both the remediation and work experience components. The remediation attendance rate seems particularly high -- across all sites, an average of 82 percent of the available hours. It is clear, however, that STEP youth responded to this summer experience in ways that are similar to their responses to regular school. Those youth who attended regularly during the school year were likely to attend STEP's remediation classes regularly; those whose school year absence rates were high were likely to miss many remediation sessions.

Implementation of the Components

Work experience, the most familiar component in the program mix, posed no major obstacles despite the challenge of finding half-time assignments and work appropriate for 14-year-olds. Jobs were for the most part typical of those generally found in SYETP programs. There was an emphasis on low-skill maintenance and clerical work, though opportunities to work with young children or in various arts projects were available in some sites.

Life skills instruction, although new to the sites, was delivered with a high degree of success. This seems due, in part, to the fact that there was a uniform curriculum that provided teachers with a lesson plan, materials and resource information. It was also due to the engaging content and format of the curriculum, both of which had been successfully piloted prior to the demonstration. In addition, teachers were given training in the use of the curriculum, and practice in dealing with difficult situations that might occur in their classrooms. Small group visits to local health clinics were well attended -- very few enrollees failed to receive their parents' consent to participate.

Remediation, the newest and least tried component of the STEP model, was inevitably the most difficult for sites to implement. Participants in fact received 90 hours of concentrated reading and math instruction, using a curriculum that was competency-based, self-paced and individualized, and that used computers to augment the instruction. Nevertheless, significant problems were encountered in the implementation of this component. These include instances of lack of clarity between P/PV and the STEP sites regarding the content and philosophy of the remediation component and inadequate resources in the sites for curriculum development, teacher training and support.

PROGRAM IMPACT

Learning Outcomes

Participation in STEP's first summer -- the first eight weeks of a 15-month treatment -- had a significant impact on the academic performance of target youth. At the end of the summer, treatment youth outscored their control group counterparts in both reading and math by approximately one-quarter of a grade equivalent. The impact of participation in STEP varied as a function of sex, race and site. It had a differential effect on boys and girls, for instance. Treatment girls scored significantly higher than control group girls on the reading post-test, while the scores of treatment and control group boys were virtually the same. In math, however, both boys and girls in the treatment group significantly outscored their control group counterparts.

Similarly, the impact of participation in STEP varied among the racial/ethnic groups in the sample. Considering the impact on reading, it is clear that Hispanics and Asians benefited most from participation in STEP, outscoring their control group counterparts by 17 points and 10 points, respectively. This translates into an improvement of half a grade equivalent (approximately) for Asians and three-quarters of a grade for Hispanics.

Hispanic youth also appear to be the group to benefit most in math: at post-test, Hispanic youth in the treatment group outscored Hispanics in the control group by 20 points or a half a grade equivalent. Black treatment youth also gained relative to their control group counterparts in math -- by a quarter of a grade.

Thus it appears that participation in STEP has been of greatest benefit to Hispanic youth, a group whose high school dropout rate -- at 45 percent -- is double the rate for blacks and almost three times the rate for whites. Asians and blacks have derived solid benefits as well.

The impact of participation in STEP also varied among the sites. In both reading and math, treatment youth in Boston scored significantly higher than control group youth at post-test. In Fresno, treatment youth scored significantly higher in math than youth in the control group, while in Seattle, treatment youth's reading scores were significantly higher than those of the control group. However, it is clear that participants in Portland -- a site which experienced considerable difficulty implementing its remediation component -- did not do as well vis-a-vis their control group counterparts as participants at other sites. In fact, they were significantly outscored by the control group.

When the relationship between academic performance and the amount of remediation a treatment youth received was examined, a significant relationship was found for reading. It indicates that a treatment youth gains 1.1 points in his post-test reading score for every hour spent in remediation classes. This is a fairly substantial impact, suggesting that had it been possible to increase by 10 hours the amount of time a given treatment youth spent in remediation classes, it would have improved his post-test reading score by 10 points.

The potential of the STEP model to achieve academic gains was demonstrated among some subsets of the treatment group. Asians in the sample achieved gains from pre- to post-test in both reading and math as did treatment youth at three of San Diego's remediation sites. These results, while encouraging, are clearly not all that could be desired. In particular it is disappointing that the impact was, for the most part, attained as a result of stemming the learning losses that would have occurred in the absence of the treatment, rather than producing actual gains for treatment youth. For this reason, a number of modifications to the remediation component will be made for next summer. They include P/PV development of curricular modules that will be adopted in a fairly uniform manner by the sites, curriculum content that will link all the program components, and an earlier development process utilizing a lead teacher.

Since, as was true of the pilot research, the positive impact on reading and math scores is dependent upon control group losses, a thorough exploration of their validity was undertaken. These efforts included building into the research design a system of incentives to increase control group motivation to perform well at post-test, and an analytic investigation of the pattern of pre-test to post-test responses of both treatment and control group youth. These analyses generally supported the notion that the control group losses reflect true declines in academic skill levels.

Life Skills Outcomes

Participation in STEP had a substantial impact upon participants' knowledge of fertility-related information. Across every site,

race and sex subgroup, STEP youth significantly increased their birth control knowledge. Moreover, there is reason to believe that participation in STEP is beginning to have an impact on actual behavior.

Treatment boys were less likely to report sexual activity during the program period than boys in the control group and it appears that these differences in the self-reported sexual behavior of treatment and control group boys reflect actual differences in behavior. STEP, then, appears to have demonstrated a positive impact on the sexual behavior of boys, an effect unusual among pregnancy prevention programs. This appears to be a direct result of the effort taken in the design of the life skills curriculum to provide a balanced intervention for both boys and girls.

In its first summer, STEP yielded an interesting and complex set of findings, on balance quite encouraging. These results from the first eight weeks of a 15-month treatment demonstrate the potential of the program model as a vehicle for facilitating positive outcomes for low-income youth at risk of dropping out of school. Drawing from the lessons learned this summer, a number of refinements to the basic model are being made. These include modifications to the remediation curriculum and the development of an expanded life skills curriculum for returning youth.

REFERENCES

- Branch, Alvia Y., Milliner, Julita, et. al., Summer Training and Education Program, Report on the Pilot Experience. Public/Private Ventures, Philadelphia, Pennsylvania: March 1985.
- Card, J.J., and Wise, L.L. "Teenage Mothers and Teenage Fathers: The Impact of Early Childbearing on the Parents' Personal and Professional Lives." Family Planning Perspectives, 1978, Vol. 10, 199-205.
- Coombs, J. and Cooley, W.W. "Dropouts: In High School and After High School." American Educational Research Journal, 1968, Vol. 5, 343-358.
- George Farkas, et. al., Final Program Impacts of the Youth Incentive Entitlement Pilot Projects. Abt Associates, Inc., Cambridge, Massachusetts: 1983.
- Heyns, Barbara, Summer Learning and the Effect of Schooling. Academic Press, New York: 1978.
- Kapsis, Robert and William Protash, Summer Motivation and Retention. City University of New York: 1983.
- Levin, Henry M., The Educationally Disadvantaged: A National Crisis. Public/Private Ventures, Philadelphia, Pennsylvania: July 1985.
- Malizio, Andrew G. and Whitney, Douglas R. Educational Credentials in Employment: A Nationwide Survey. Paper presented at the Lifelong Learning Conference, College Park, Maryland: February 16-17, 1984.
- Presser, H.B. "Social Consequences of Teenage Childbearing." In Chilman, c. (ed.) Teenage Childbearing: Recent Research on the Determinants and Consequences. Government Printing Office, Washington, D.C.: 1978.
- Sum, Andrew, et. al., Educational Attainment, Academic Ability, and the Employability and Earnings of Young Persons: Implications for the Planning and Design of JTPA Youth Programs. Prepared for a Symposium on Serving Disadvantaged School Dropouts. Brandeis University, December, 1983.
- Taggart, Robert, A Fisherman's Guide: An Assessment of Training and Remediation Strategies. W.E. Upjohn Institute, Kalamazoo, Michigan: 1981.
- Turnage, Wayne, A. Factors Influencing High School Dropouts: A Literature Review. Manpower Demonstration Research Corporation, New York, NY: 1982.

APPENDIX A
TEST FLOOR EFFECTS

Appendix A

Test Floor Effects

During the pilot program, control group losses in both reading and math test scores were smaller than the corresponding control group losses in the first summer of actual demonstration. Also, the apparent gains for the participants in the pilot were not realized in the demonstration. Part of the divergence in the impact of the pilot and demonstration projects may be attributable to the variation in the program among sites. Boston was the only site with both a pilot and demonstration.

In Boston the relative impact of the program -- how much the treatment group's score changed during the summer compared to the change for the control group net of other facts -- was very nearly the same for the demonstration as it was for the pilot (30.10 versus 22.47 for reading and 42.90 versus 39.46 in math respectively), even though the absolute outcome was much lower for both groups this year.

The apparent divergence in absolute impacts may, in fact, simply result from the use of different levels of the MAT test in the pilot and demonstration programs. The Advanced Level I MAT was used in the pilot while the Intermediate test was used in the demonstration. Test floor effects may have resulted in an under-estimate of control group summer losses and an over-estimate of treatment gains in the pilot program.

The pilot experience showed that the Advanced I Level of the MAT was too advanced for the youth who were participating in the STEP program. The Advanced I Level of the test is aimed at the 7.0 to 9.9 average grade level of achievement. STEP participants, however, tended to score close to the 5.0 grade equivalent in reading and about the 7.0 level in math. The next lower level of the MAT, the Intermediate Level, which was designed for the 5.0 to 6.9 range, was chosen for the demonstration since it should more accurately measure the reading ability of STEP kids and perform equally well in measuring the math ability. The following discussion explores the biases introduced by test floor effects and their implications for differences between the pilot and demonstration results.

On any multiple choice test, random selection of answers will result in some correct answers. The test floor is defined as the expected number of correct answers resulting from random guessing. For the Advanced I, that represents 14 correct items on the reading test for an average scaled score of 623, and 13 correct answers on the math test for an average scaled score of 597, when considering both versions of the same test. For the Intermediate level of the MAT the floor is exactly the same number of items

correct as for the Advanced I math test, since they are of similar length and both have four choices per question. On the Intermediate reading test, there are 60 possible answers, with a floor of 15, corresponding to an average scaled score of 595. However, since the Intermediate is targetted at a lower level of achievement, the scaled score that this floor is equivalent to is much lower -- an average of 510 considering both versions of test.

If a person's ability is below the floor at the Intermediate Level MAT, then it is possible that the individual would obtain higher scaled scores on the Advanced test, since, on average, the individual would obtain the test floor in both tests. The test floor, however, is associated with a higher grade level in the Advanced test. Thus, the net effect of using a test in which a significant number of youth are performing at the floor level is to overstate the scaled scores of the entire group.

Test floor effects are important for understanding the absolute differences in the pilot and between the pilot and demonstration results. In order to investigate this issue, two simplifying assumptions are necessary. First, no one scores below the floor. While this is not always true, it will be true on average. Second, STEP doesn't help anyone whose ability is below the floor of the Intermediate test. This seems reasonable since STEP was never intended to be able to help a 14-year-old who reads below the grade level associated with the floor at the Intermediate test.

The implications of the previous assumptions are as follows: (1) if a youth had an ability level which is below the floor of the Intermediate Level of the test at the beginning of the summer, then neither level of the test would measure his decline during the summer, (2) if, however, a youth was at or below the floor of the Advanced I and above the floor of the Intermediate, then his loss would be measured only by the Intermediate while the Advanced I would show no change, (3) if a youth just slightly above the Advanced I floor suffered a loss which would put him under the floor, this would show up as a larger loss on the Intermediate Level of the test. Since under these conditions all gains would be recorded by either version of the test, while some losses are either not measured or are gauged as smaller on the Advanced I version, the overall average change in scores over the summer when evaluated by the Advanced I Level would indicate either a smaller loss or a larger gain than would the Intermediate Level.

In order to test this hypothesis the floor that would have prevailed, had this summer's sample been given the Advanced I version of the test, was artificially imposed by resetting the scaled score of anyone who scored below 623 on reading or 597 in math to be equal to 623 or 597 on their respective test. This was done for both the pre-test and post-test. Then the change in

test performance over the summer was recalculated for all those youth who took both tests. The results are represented in Table A-1 along with the results for this summer's actual observed outcome.

The table indicates that there were enough youth at the test floor so that there would have been smaller measured losses had the Advanced test been given to the demonstration youth. (Except in the case of treatment group math scores, which would have changed from losses to gains.)

It should be noted that this exercise still does not generate the magnitude of gains for the treatments that was posted during the pilot project, which means that while it may explain some of the differences, it is not a complete explanation.

Table A-1
Pre- to Post-test Score Change
With and Without Floor Effect

	<u>CHANGE ACTUALLY OBSERVED</u>	<u>CHANGE WITH SIMULATED FLOOR</u>
<u>Treatments</u>		
Reading	-8.6	-6.7
Math	0.0	0.0
<u>Controls</u>		
Reading	-14.3	-12.5
Math	-16.5	-10.9

APPENDIX B

ATTRITION

Appendix B

Attrition

In any complex demonstration like STEP, there will be sample attrition. The consequences of attrition are most serious if the treatment program directly affects the pattern of attrition. If, for example, the attrition rate for blacks in the treatment group were markedly different from those in the control group, then it is clear that the treatment is causing a change in the composition of the sample. This can introduce sample selection bias into the estimates of programmatic impacts. On the other hand, the existence of differences in treatment and control group characteristics is not a problem since this issue is easily handled using multivariate analytical techniques, rather it is only treatment-induced differences between treatment and control samples which are of concern.

There are 1,146 youth who attended both testing sessions and provided information on all items which went into constructing the variables of interest in the basic regression models. The primary demographic characteristics of the sample used to estimate the learning impact of the program are very similar for the treatment and control groups. Chi-squared statistics obtained when each of the demographic characteristics -- sex, race, and age -- are taken, one at a time, along with the presence or absence of treatment, indicate no significant differences in means between the control and treatment groups. Since the original distributionss across treatments and controls is random, one may conclude that attrition has not caused selection bias in this sample.

An analogous exploration of the distribution of the demographic characteristics was carried out for the Life Skills (LSO) sample. It is apparent that for the LSO analyses the racial/ethnic distribution of the treatment and control group differed somewhat. Note the LSO sample is smaller because of many more missing observations for important variables. The non-random distribution of races does not necessarily imply the presence of a selectivity problem, but it does admit the possibility.

To investigate any potential selectivity problem due to attrition, it was necessary to determine if the treatment program was a determinant of being in the LSO sample. A probit model estimating the probability of being in the LSO sample, using age, race, sex and treatment as explanatory variables, was estimated. In this model, an estimated coefficient for treatment which is significantly different from zero would indicate that, controlling for demographic differences, there is something about (or at least, correlated with) the treatment which affects likelihood of being in the sample, and selectivity bias would be a major concern.

The results of this investigation indicate (Table B-1) that there is no need for any adjustments or corrective action to compensate for sample attrition. This is because the coefficient of the treatment variable is not significantly different from zero.

This same technique was also used to test if the treatment status of the participants affected observed patterns of attrition. The estimated test statistic indicated that there are no significant differences in attrition between the treatment and control groups. These results are shown in Table B-2.

Since there is no evidence of selectivity bias, the impacts which have been estimated in this report represent unbiased and consistent estimates of the treatment effect on the sample for which there is complete data.

Table B-1

Probit Estimation for Learning Sample Membership

Explanatory Variable	Coefficient
Constant term	-0.51 (4.53)**
Male	-0.20 (-2.89)**
Hispanic	0.10 (0.80)
Black	0.17 (1.59)
Asian	0.35 (2.20)**
Age	0.01 (0.90)
Treatment	0.02 (0.28)

Value of Likelihood Function for Restricted Model = 947.36
 Value of Likelihood function for Unrestricted Model = 939.36
 Chi-squared = 16,000* with 6 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

Table B-2

Probit Estimation for Life Skills Sample Membership

Explanatory Variable	Coefficient
Constant term	0.32 (2.96)**
Male	-0.03 (-4.69)**
Hispanic	0.00 (0.01)
Black	0.10 (0.94)
Asian	0.23 (1.80)*
Age	-0.03 (0.43)
Treatment	-0.03 (-0.49)

Value of Likelihood Function for Restricted Model = 1080.30
 Value of Likelihood function for Unrestricted Model = 1066.30
 Chi-squared = 27.40* with 6 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

APPENDIX C
CONTROL GROUP LOSSES

Appendix C

Control Group Losses

Since much of the net impact on learning of the STEP program depends on the magnitude of the control group losses, it is important to determine whether these losses are real or spurious. The results could be spurious if control and treatment group members were not equally motivated at post-test. Treatment youth may have some incentive to show what they have learned in their weeks of remediation (spurred on by the enthusiasm of the teachers) while the control youth have been out of a school setting for three months and probably lack such motivation. Thus, the issue of equally motivating the treatment and control youth on the follow-up test was a major concern in the design of the STEP demonstration.

An incentive scheme needed to be developed to motivate both the treatment and control group youth. Such an incentive scheme needed to be large enough to motivate youth, yet not be prohibitively expensive. We ruled out rewarding youth with a certain amount of money (x cents) for each right answer, since the expense associated with this scheme was excessive. Instead, a lottery in which the chance of winning increased with test performance was devised.

To provide equal incentive for treatment and control students to do their best on each test, a lottery ran concurrently with the pre- and post-tests at each site. Prior to taking the tests, the students were informed that they would receive one chance to win for taking the test and one chance for each correct answer on their tests. Fifteen winners of \$50 each were then chosen from each group at each site, so that approximately one youth out of 10 received a prize at each administration of the test. Obviously, one's likelihood of winning is greatly increased by how well one performs on the tests.

Still, it is necessary to assess whether differential motivation between the two groups might be responsible for control group losses, despite the incentive scheme. One method involved dividing the test in half and comparing the number of correct answers on the first half with the number correct on the last half. Without making assumptions about reduced concentration towards the end of the test, or the tendency to omit or randomly guess, differential motivation could show up as a divergent pattern between treatment and control groups in the first-half to last-half correct answers. The results presented in Table C-1 indicate remarkably similar patterns, both between treatments and controls, as well as between pre-test and post-test.

Table C-1

Half Test Score Comparisons

<u>TEST</u>	<u>TREATMENT</u>	<u>CONTROL</u>
Reading pre-test	22.8 - 16.9	22.7 - 17.3
Reading post-test	21.3 - 16.9	21.1 - 16.4
Math pre-test	15.5 - 16.7	15.9 - 16.9
Math post-test	15.4 - 16.6	15.1 - 15.7

The second approach was to estimate the reliability for each test -- separately for treatment and controls. Reliability in educational measurement means the consistency with which a set of test scores measure whatever they are intended to measure. This is usually operationally defined as the coefficient of correlation between that set of scores and another set of scores on an equivalent test obtained independently from the members of the same group. A test for which one would always obtain the same average score from the same group of kids would receive a computed reliability of 1.00. This assumes that the youth do not learn from the test itself.

In Table C-2 the estimated reliability coefficients for pre- and post-tests (math and reading) for the treatments and controls are presented. These were estimated using Kuder-Richardson Formula 20. The similarity between the treatment and control groups certainly seems to indicate that whatever the level of motivation was, it was the same for the two groups. This same finding holds true even when examining only the Boston youth of the STEP program, where the extremely poor performance of the controls might make their motivation particularly suspect. Again the Kuder-Richardson Formula 20 was computed and the results compared between treatments and control on both pre-test and post-test. These coefficients are presented in Table C-3 and show roughly comparable reliability coefficients for treatment and control youth.

Therefore, combining this evidence with that recorded above and the overall similarity of the distribution of test scores presented elsewhere in this report, it is concluded that the measured impacts of the STEP program (and the control group losses) are not biased by differential lack of motivation to perform on the part of the controls.

Table C-2
Achievement Test Reliability Measurements

<u>TEST</u>	<u>TREATMENT</u>	<u>CONTROL</u>
Reading pre-test	.932	.928
Reading post-test	.948	.944
Math pre-test	.880	.868
Math post-test	.908	.902

Table C-3
Achievement Test Reliability
Measurements for Boston Only

<u>TEST</u>	<u>TREATMENT</u>	<u>CONTROL</u>
Reading pre-test	.937	.923
Reading post-test	.943	.937
Math pre-test	.850	.814
Math post-test	.882	.902

APPENDIX D
TECHNICAL DATA TABLES

Table D-1

STEP Analysis Variable Descriptions

<u>VARIABLE</u>	<u>DESCRIPTION</u>	<u>TYPE</u>
Absent much during past school year	Participant was absent 11 or more days during the 1984-85 school year.	Dummy (0,1)
Age	Self-explanatory.	Dummy (0,1)
Asian	Self-explanatory.	Dummy (0,1)
Woman responsible for birth control (base)	Participant responded that only the woman is responsible for birth control (reported at pre-test).	Dummy (0,1)
Woman responsible for birth control (end)	As above (reported at post-test).	Dummy (0,1)
Birth control availability score (base)	Participant's knowledge of the availability of birth control (index based on number of correct answers to eight availability questions on the pre-test).	Index
Birth control availability score (end)	As above (reported at post-test).	Index
Outcome of teen motherhood (base)	Participant understands the consequences of teenage pregnancy to the mother, i.e., she will probably have to leave school permanently (reported at pre-test).	Dummy (0,1)
Outcome of teen motherhood (end)	As above (reported at post-test).	Index
Best age to have children (base)	Participant's belief of the best age to have first child (reported at pre-test).	Dummy (0,1)
Best age to have children (end)	As above (reported at post-test).	Dummy (0,1)
Black	Self-explanatory.	Dummy (0,1)
Boston	Self-explanatory.	Dummy (0,1)

Table D-1 (cont'd)

<u>VARIABLE</u>	<u>DESCRIPTION</u>	<u>TYPE</u>
Position in birth order	Participant's birth order (e.g., would be three if participant was the third child born in the family).	Categorical
Contraceptive use meaning 1 (base) (CONTRA 1)	Participant used birth control at last sexual contact, given that the participant has had sex within the past two months (reported at pre-test).	Dummy (0,1)
Contraceptive use meaning 1 (end)	As above (reported at post-test).	Dummy (0,1)
Contraceptive use meaning 2 (base)	See variable CONTRA 1 above, given that the participant has ever used contraception (reported at pre-test).	Dummy (0,1)
Contraceptive use meaning 2 (end)	As above (reported at post-test).	Dummy (0,1)
Ever pregnant	Participant was or is currently pregnant.	Dummy (0,1)
No-show or dropped from program	Participant left program before completion or never participated in the program.	Dummy (0,1)
Enjoys math (base)	Measure of math enjoyment from 1 (very much) to 5 (not at all). (Reported at pre-test.)	Categorical
Enjoys reading (base)	Measure of reading enjoyment from 1 (very much) to 5 (not at all). (Reported at pre-test.)	Categorical
Expects post high school education (base)	Participant expects to obtain at least some post-high school education (reported at pre-test).	Categorical
Ever had sex	Self-explanatory.	Dummy (0,1)

Table D-1 (cont'd)

<u>VARIABLE</u>	<u>DESCRIPTION</u>	<u>TYPE</u>
Family size	Number of members of household including participant.	Categorical
Age of first sex	Self-explanatory.	Continuous
Fresno	Self-explanatory.	Dummy (0,1)
Grade	Self-explanatory.	Categorical
Hispanic	Self-explanatory.	Dummy (0,1)
Know consequences of pregnancy (base)	Two, if participant knows consequences of teenage pregnancy to mother (see above) and the cost of raising a child to age 18 (see below). One, if participant knows either. Zero, if participant knows neither. (Reported at pre-test.)	Index
Know consequences of pregnancy (end)	As above (reported at post-test).	Index
Knowledge of birth control score (base)	Participant's knowledge of birth control methods (index based on number of correct answers to 12 birth control usage questions on the pre-test).	Index
Knowledge of birth control score (end)	As above (reported at post-test).	Index
Problem with English	A language in addition to English is spoken at home and the participant indicated problems speaking or understanding English.	Dummy (0,1)
Male	Self-explanatory.	Dummy (0,1)
Change in math score	Change observed in participant's score from pre-test math test to post-test math test.	Continuous

Table D-1 (cont'd)

<u>VARIABLE</u>	<u>DESCRIPTION</u>	<u>TYPE</u>
Mexican-American	Self-explanatory.	Dummy (0,1)
Lives without father	Lives without father in the household.	Dummy (0,1)
Did not participate in program	Self-explanatory.	Dummy (0,1)
Number jobs (base)	Number of jobs participant thinks a woman can do (reported at pre-test).	Index
Number jobs (end)	As above (reported at post-test).	Index
Other	Self-explanatory.	Dummy (0,1)
Portland	Self-explanatory.	Dummy (0,1)
Receives AFDC and/or foodstamps	Self-explanatory.	Dummy (0,1)
Puerto Rican	Self-explanatory.	Dummy (0,1)
Change in reading score	Change observed in participant's score from pre-test reading to post-test reading test.	Continuous
Sex during previous two months (base)	Self-explanatory.	Dummy (0,1)
Sex during previous two months (end)	Self-explanatory.	Dummy (0,1)
Repeated a grade	Self-explanatory.	Dummy (0,1)
San Diego	Self-explanatory.	Dummy (0,1)
Seattle	Self-explanatory.	Dummy (0,1)
Previously had sex education course	Self-explanatory.	Dummy (0,1)
Child-raising cost (base)	Participant knows the approximate cost of raising a child to age 18 (i.e., over \$75,000). (Reported at pre-test.)	Dummy (0,1)

Table D-1 (cont'd)

<u>VARIABLE</u>	<u>DESCRIPTION</u>	<u>TYPE</u>
Child-raising cost (end)	As above (reported at post-test).	Dummy (0,1)
Math pre-test	Scaled score from math portion of MAT survey test, intermediate level (pre-test).	Continuous
Math post-test	As above (reported at post-test).	Continuous
Reading pre-test	Scaled score for reading portion of MAT survey test, intermediate level (pre-test).	Continuous
Reading post-test	As above (reported at post-test).	Continuous
Total program hours	Self-explanatory.	Continuous
Total remediation hours	Total remediation and life skills hours.	Continuous
Treatment/Asian interaction term	Self-explanatory.	Dummy (0,1)
Treatment/Black interaction term	Self-explanatory.	Dummy (0,1)
Treatment/Boston interaction term	Self-explanatory.	Dummy (0,1)
Treatment/Fresno interaction term	Self-explanatory.	Dummy (0,1)
Treatment/Hispanic interaction term	Self-explanatory.	Dummy (0,1)
Treatment/Male interaction term	Self-explanatory.	Dummy (0,1)
Treatment/Portland interaction term	Self-explanatory.	Dummy (0,1)
Treatment/2nd reading quartile interaction	Self-explanatory.	Dummy (0,1)
Treatment/3rd reading quartile interaction	Self-explanatory.	Dummy (0,1)

Table D-1 (cont'd)

<u>VARIABLE</u>	<u>DESCRIPTION</u>	<u>TYPE</u>
Treatment/4th reading quartile interaction	Self-explanatory.	Dummy (0,1)
Treatment/San Diego interaction term	Self-explanatory.	Dummy (0,1)
Treatment/reading pre- test interaction term	Self-explanatory.	Dummy (0,1)

Table D-2

Descriptive Statistics for Treatment and Control Group

Explanatory Variable	TREATMENT		CONTROL	
	Number Non-Missing	Mean (Standard Deviation)	Number Non-Missing	Mean (Standard Deviation)
<u>Educational Characteristics</u>				
Enjoys reading (base)	748	2.71 (1.05)	737	2.76 (1.04)
Enjoys math (base)	749	2.51 (1.16)	739	2.56 (1.19)
Absent much in past school year	738	.25	731	.30
Expects post-high school education (base)	728	.72	728	.70
Repeated a grade	698	.28	709	.28
Grade	747	8.51 (.74)	734	8.49 (.74)
<u>Demographic Characteristics</u>				
Male	798	.53	795	.53
Boston	798	.21	795	.21
Fresno	798	.20	795	.20
Portland	798	.19	795	.20
San Diego	798	.20	795	.19
Seattle	798	.21	795	.20
Age	798	.50	795	.49
Problem with English	798	.09	795	.08

Table D-2 (cont'd)

Explanatory Variable	TREATMENT		CONTROL	
	Number Non-Missing	Mean (Standard Deviation)	Number Non-Missing	Mean (Standard Deviation)
Family size	733	5.23 (2.08)	724	5.24 (2.07)
Receives AFDC and/or foodstamps	796	.62	790	.60
Lives without father	725	.51	710	.49
Position in birth order	725	2.96 (2.30)	743	2.90 (2.34)
Asian	798	.14	795	.14
Black	798	.61	795	.57
Other	798	.10	795	.13
Hispanic	798	.16	795	.16
Number jobs (base)	745	5.84 (2.63)	734	5.62 (2.71)
Number jobs (end)	680	6.39 (2.56)	664	6.04 (2.63)
No-show or dropped from program	796	.19	795	.17
Puerto Rican	124	.17	129	.17
Mexican-American	124	.46	128	.50
Sexual Characteristics and Sexual Question Responses				
Age of first sex	335	12.01 (1.94)	330	11.94 (1.96)
Ever had sex	672	.52	669	.51
Ever pregnant	100	.10	92	.13

Table D-2 (cont'd)

Explanatory Variable	TREATMENT		CONTROL	
	Number Non-Missing	Mean (Standard Deviation)	Number Non-Missing	Mean (Standard Deviation)
Previously had sex education course	619	.62	621	.57
Outcome of teen motherhood (base)	743	.74	734	.71
Outcome of teen motherhood (end)	567	.96	499	.93
Child-raising cost (base)	740	.32	731	.36
Child-raising cost (end)	680	.41	661	.41
Know consequences of pregnancy (base)	739	1.06 (.66)	729	1.08 (.68)
Know consequences of pregnancy (end)	565	1.37 (.55)	498	1.34 (.56)
Woman responsible for birth control (base)	739	.24	722	.24
Woman responsible for birth control (end)	672	.11	660	.20
Sex during previous two months (base)	656	.28	654	.31
Sex during previous two months (end)	646	.32	627	.34
Contraceptive use meaning 2 (base)	100	.70	99	.77
Contraceptive use meaning 2 (end)	114	.70	99	.77
Contraceptive use meaning 1 (base)	174	.40	185	.41

Table D-2 (cont'd)

Explanatory Variable	TREATMENT		CONTROL	
	Number Non-Missing	Mean (Standard Deviation)	Number Non-Missing	Mean (Standard Deviation)
Contraceptive use meaning 1 (end)	192	.46	200	.42
Birth control availability score (base)	751	2.99 (1.94)	743	3.06 (1.94)
Birth control availability score (end)	687	4.88 (1.96)	668	3.25 (1.88)
Knowledge of birth control score (base)	751	4.76 (2.38)	743	4.92 (2.36)
Knowledge of birth control score (end)	687	7.25 (2.68)	668	5.08 (2.48)
Best age to have children (base)	663	22.45 (4.72)	667	22.36 (4.18)
Best age to have children (end)	638	22.04 (3.88)	619	22.42 (4.44)
<u>STEP Test Data</u>				
Reading pre-test	752	715.98 (59.09)	743	718.08 (56.77)
Math pre-test	686	687.07 (80.38)	742	693.64 (76.13)
Reading post-test	686	709.02 (68.60)	666	704.35 (63.78)
Math post-test	686	689.38 (92.28)	665	679.09 (91.10)
Change in reading	673	-8.55 (53.27)	649	-14.32 (48.00)
Change in math	672	.04 (70.20)	648	-16.47 (65.76)
Total remediation hours	788	87.59 (36.05)	795	.65 (7.39)
Total program hours	787	168.06 (69.66)	790	167.91 (70.89)

Table D-3

Descriptive Statistics Data Used in Learning Analysis
Treatment and Control Group

<u>Explanatory Variable</u>	<u>Treatments</u>	<u>Controls</u>
<u>Educational Characteristics</u>		
Repeated a grade	.28	.26
Grade last year	8.48 (.73)	8.48 (.74)
Absent much during past school year	.24	.24
Expects post-high school education (base)	.73	.71
Enjoys reading (base)	2.71 (1.04)	2.75 (1.03)
<u>Demographic Characteristics</u>		
Male	.50	.51
Black	.61	.57
Hispanic	.15	.15
Asian	.15	.15
Receives AFDC and/or foodstamps	.65	.62
Family size	5.27 (2.03)	5.28 (2.08)
Lives without father	.49	.48
Position in birth order	3.03 (2.42)	2.88 (2.37)
Problems with English	.10	.08
Boston	.21	.21
Fresno	.21	.18
Portland	.19	.19
San Diego	.18	.21

Table D-3 (cont'd)

Descriptive Statistics Data Used in Learning Analysis
Treatment and Control Group

<u>Explanatory Variable</u>	<u>Treatments</u>	<u>Controls</u>
<u>STEP Test Data</u>		
Reading pre-test	718.99 (57.89)	720.66 (56.36)
Reading post-test	711.53 (67.38)	706.80 (64.41)
Math pre-test	692.36 (80.61)	698.72 (74.77)
Math post-test	692.20 (89.22)	682.98 (92.82)
Sample size	576	568

Table D-4

Descriptive Statistics Data Used in Life Skills
Treatment and Control Group

<u>Explanatory Variable</u>	<u>Treatments</u>	<u>Controls</u>
<u>Educational Characteristics</u>		
Absent much during past school year	.24	.27
Expects post-high school education (base)	.74	.73
<u>Demographic Characteristics</u>		
Age	.51	.47
Male	.46	.50
Black	.61	.58
Hispanic	.15	.15
Asian	.16	.14
Receives AFDC and/or foodstamps	.66	.63
Family size	5.31 (1.99)	5.29 (2.12)
Lives without father	.48	.48
Position in birth order	2.91 (2.30)	2.89 (2.35)
Problem with English	.10	.07
Boston	.19	.21
Fresno	.20	.18
Portland	.21	.21
San Diego	.20	.21
<u>Sexual Characteristics and Sexual Question Responses</u>		
Ever had sex	.48	.47
Previously had sex education course	.63	.60

Table D-4 (cont'd)

<u>Explanatory Variable</u>	<u>Treatments</u>	<u>Controls</u>
Birth control availability score	3.23 (1.94)	3.19 (1.91)
Knowledge of birth control score	4.96 (2.32)	5.02 (2.32)
Best age to have children	22.47 (3.88)	22.53 (4.09)
Woman responsible for birth control (base)	.25	.22
Know consequences of pregnancy (base)	1.09 (.66)	1.10 (.69)
<u>STEP Test Data</u>		
Reading pretest	726.51 (56.65)	724.56 (58.23)
Reading posttest	718.48 (66.69)	709.85 (63.81)
Math pretest	701.08 (77.41)	702.72 (74.54)
Math posttest	702.37 (87.75)	686.57 (92.63)
Sample size	464	471

Table D-5

Reading and Math Regressions
Pooled Data From All Sites

Explanatory Variable	READING EQUATIONS		MATH EQUATIONS	
	Basic	With Interactions	Basic	With Interactions
Constant term	128.80 (4.91)**	129.27 (-4.90)**	50.65 (1.47)	61.02 (1.75*
Treatment	5.68 (2.00)**	19.78 (1.72)*	12.64 (3.33)**	-1.10 (-0.07)
Treatment/Male Interaction		-10.97 (1.94)*		-0.13 (-0.02)
Treatment/Black Interaction		-9.10 (-0.94)		5.03 (0.39)
Treatment/Hispanic Interaction		10.53 (0.87)		11.14 (0.69)
Treatment/Asian Interaction		6.93 (0.57)		17.38 (1.07)
Treatment/Boston Interaction		13.35 (1.47)		37.40 (3.08)**
Treatment/Fresno Interaction		-10.22 (-1.04)		13.50 (1.03)
Treatment/Portland Interaction		-21.13 (-2.32)**		-10.91 (-0.89)
Treatment/San Diego Interaction		-12.82 (-1.43)		-9.99 (-0.83)
Male	-8.97 (-3.11)**	-2.85 (-0.71)	-16.17 (-4.22)**	-14.87 (-2.75)**
Black	-4.07 (-0.84)	-0.76 (-0.12)	-9.91 (-1.53)	-12.43 (-1.48)

Table D-5 (cont'd)

Explanatory Variable	READING EQUATIONS		MATH EQUATIONS	
	Basic	With Interactions	Basic	With Interactions
Asian	18.83 (2.71)**	13.76 (1.57)	27.54 (2.91)**	18.73 (1.59)
Receives AFDC and/or Food Stamps	-0.71 (-0.23)	-1.45 (-0.47)	0.60 (0.15)	-0.23 (-0.06)
Family size	-0.49 (-0.64)	-0.27 (-0.35)	-0.86 (-0.83)	-0.69 (-0.67)
Lives without father	4.61 (1.47)	5.06 (1.61)	2.59 (0.62)	3.00 (0.86)
Position in birth order	-0.90 (-1.50)	-0.92 (-1.55)	0.36 (0.45)	
Problem with English	-4.95 (-0.81)	-4.64 (-0.76)	6.31 (0.77)	
Grade last year	7.02 (3.17)**	7.24 (3.30)**	10.86 (3.65)**	11.14 (3.77)**
Repeated a grade	-4.59 (-1.28)	-5.25 (-1.48)	3.23 (0.67)	2.74 (0.57)
Much absence in past school year	-1.28 (-0.38)	-0.92 (-0.27)	4.54 (1.01)	5.47 (1.22)
Expects post-high school education (base)	-0.58 (-0.17)	-0.83 (-0.25)	2.73 (0.61)	3.17 (0.71)
Reading pre-test	0.73 (27.00)**	0.72 (26.67)**	0.23 (5.30)**	0.23 (5.23)**
Enjoys reading (base)	-2.72 (-1.90)*	-3.01 (-2.11)**		
Boston	10.74 (2.24)**	4.34 (0.67)	2.72 (0.42)	17.21 (1.97)**
Fresno	16.13 (3.24)**	21.32 (3.02)**	5.93 (0.88)	

Table D-5 (cont'd)

Explanatory Variable	READING EQUATIONS		MATH EQUATIONS	
	Basic	With Interactions	Basic	With Interactions
Portland	15.60 (3.39)**	26.06 (4.01)**	11.56 (1.87)*	
San Diego	6.94 (1.50)	14.04 (2.20)**	10.78 (1.68)*	16.44 (1.89)*
Math pre-test			0.58 (16.64)**	0.57 (16.48)**
Enjoys math (base)			-10.82 (-6.09)**	-11.25 (-6.35)**
Sample size	1145	1145	1145	1145
F-test	53.45	39.82	58.78	43.89
R-squared	0.49	0.50	0.52	0.53
Adjusted R-squared	0.48	0.49	0.51	0.52

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-5a
Reading Impacts by Site

Explanatory Variable	ESTIMATED COEFFICIENT				
	Boston	Fresno	Portland	San Diego	Seattle
Constant term	133.75 (1.88)*	189.29 (4.10)**	177.63 (2.32)**	197.63 (3.02)**	50.75 (0.98)
Treatment	22.47 (3.20)**	4.79 (1.00)	-12.43 (-1.63)	5.42 (0.81)	9.39 (1.58)
Male	-8.17 (-1.15)	-2.36 (-0.48)	-7.85 (-0.99)	-12.32 (-1.74)*	-7.73 (-1.24)
Black	2.00 (0.21)	-3.59 (-0.35)	-13.56 (-1.41)	-26.32 (-1.32)	12.99 (1.20)
Hispanic	-10.05 (-0.78)	-6.52 (-0.64)		-11.26 (-0.53)	
Asian			33.40 (1.85)*	-9.49 (-0.44)	28.47 (2.29)**
Receives AFDC and/or foodstamps	-0.63 (-0.08)	1.76 (0.31)	2.56 (0.32)	-2.72 (-0.33)	-5.06 (-0.81)
Family size	-0.02 (-0.0*0)	-1.70 (-1.43)	0.36 (0.18)	1.38 (0.74)	-2.15 (-1.31)
Lives without father	14.30 (1.91)*	-1.42 (-0.27)	5.73 (0.72)	2.14 (0.28)	3.90 (0.56)
Position in birth order	-0.25 (-0.17)	-0.05 (-0.05)	0.17 (0.12)	-2.99 (-2.05)**	-1.50 (-0.98)
Problem with English		-18.57 (-1.35)	-18.80 (-0.94)	10.48 (0.98)	-3.27 (0.30)
Grade last year	7.97 (1.07)	4.88 (1.24)	7.04 (1.17)	5.54 (1.13)	7.76 (1.97)*
Repeated a grade	-11.82 (-1.47)	0.12 (0.02)	4.77 (0.46)	-11.30 (-1.29)	-11.81 (1.51)
Much absence in past school year	-9.49 (-1.14)	3.43 (0.68)	5.24 (0.63)	-14.53 (-1.47)	4.35 (0.60)

Table D-5a (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT				
	Boston	Fresno	Portland	San Diego	Seattle
Expects post-high school education (base)	-2.05 (-0.29)	-4.10 (-0.70)	12.99 (1.45)	-10.98 (-1.21)	2.04 (0.29)
Reading pre-test	0.69 (10.46)**	0.71 (16.63)**	0.67 (8.63)**	0.72 (10.87)**	0.83 (14.28)**
Enjoys reading (base)	1.87 (0.57)	-3.48 (-1.33)	-2.80 (-0.68)	-8.31 (-2.39)**	-3.53 (-1.21)
Sample size	241	223	221	223	237
F-test	12.99	25.07	8.84	12.83	21.27
R-squared	0.45	0.65	0.39	0.46	0.56
Adjusted R-squared	0.41	0.62	0.35	0.46	0.56

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-5b

Math Impact by Site

Explanatory Variable	ESTIMATED COEFFICIENT				
	Boston	Fresno	Portland	San Diego	Seattle
Constant term	27.92 (0.31)	65.60 (0.95)	254.96 (2.61)**	52.63 (0.63)	8.30 (0.12)
Treatment	39.46 (4.60)**	19.97 (2.68)**	-5.53 (-0.56)	0.13 (0.02)	5.52 (0.65)
Male	-15.87 (-1.84)*	-24.05 (-3.10)**	-7.72 (-0.75)	-19.94 (-2.27)**	-13.52 (-1.53)
Black	-12.48 (-1.09)	-22.34 (-1.41)	-10.35 (-0.82)	-20.34 (-0.81)	3.01 (0.19)
Hispanic	-33.17 (-2.10)**	-13.05 (-0.82)		-0.59 (-0.02)	
Asian			42.39 (1.76)*	28.27 (1.00)	29.97 (1.68)*
Receives AFDC and/or foodstamps	-6.57 (-0.70)	5.18 (0.59)	-2.38 (-0.23)	-4.06 (-0.39)	6.58 (0.76)
Family size	1.08 (0.43)	-1.98 (-1.06)	-3.34 (-1.28)	0.00 (0.00)	-0.18 (-0.08)
Lives without father	9.55 (1.05)	2.20 (0.27)	3.37 (0.32)	7.65 (0.77)	1.36 (0.14)
Position in birth order	2.39 (1.32)	-0.76 (-0.52)	0.62 (0.32)	-1.72 (-0.92)	0.16 (0.07)
Problem with English		28.51 (1.32)	-23.61 (-0.89)	18.16 (1.32)	5.56 (0.36)
Grade last year	2.86 (0.32)	7.57 (1.20)	7.17 (0.89)	18.98 (3.02)**	12.49 (2.23)**
Repeated a grade	13.51 (1.40)	3.25 (0.33)	-5.74 (-0.42)	-9.99 (-0.88)	-5.73 (-0.52)
Much absence in past school year	-8.55 (-0.85)	10.47 (1.32)	11.89 (1.07)	4.61 (0.37)	9.06 (0.88)

Table D-5b (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT				
	Boston	Fresno	Portland	San Diego	Seattle
Expects post-high school education (base)	-9.09 (-1.04)	2.19 (0.24)	27.01 (2.28)**	-8.12 (-0.70)	14.87 (1.50)
Reading pre-test	0.22 (2.22)**	0.21 (2.58)**	0.08 (0.65)	0.43 (4.40)**	0.15 (1.47)
Math pre-test	0.69 (8.90)**	0.62 (9.70)**	0.49 (5.42)**	0.37 (4.19)**	0.68 (8.75)**
Enjoys math (base)	-8.90 (-2.51)**	-4.22 (-1.14)	-7.37 (-1.54)	-24.38 (-5.61)**	-11.38 (-2.85)**
Sample size	241	224	219	223	238
F-test	15.48	19.46	7.43	17.17	19.86
R-squared	0.51	0.60	0.37	0.59	0.59
Adjusted R-squared	0.48	0.57	0.32	0.55	0.56

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-5c
Reading Impact by Race

Explanatory Variable	ESTIMATED COEFFICIENT			
	Blacks	Hispanic	Asian	Whites & Others
Constant term	120.73 (3.31)**	216.80 (3.66)**	165.04 (3.04)**	63.96 (0.61)
Treatment	0.43 (0.11)	17.18 (2.46)**	10.19 (1.67)*	11.90 (1.09)
Male	-9.26 (-2.42)**	-15.10 (-2.08)**	-6.40 (-0.95)	-4.17 (-0.37)
Receives AFDC and/or foodstamps	-5.29 (-1.29)	2.72 (0.35)	-1.70 (-0.23)	22.17 (1.88)*
Family size	0.53 (0.47)	-1.37 (-0.70)	-2.43 (-1.87)*	-2.14 (-0.70)
Lives without father	6.01 (1.48)	10.23 (1.28)	-6.65 (-0.84)	-0.46 (-0.04)
Position in birth order	-1.02 (-1.34)	-2.69 (-1.75)*	0.77 (0.53)	1.05 (0.38)
Problem with English		-17.61 (-1.46)	-3.89 (-0.61)	
Grade last year	6.96 (2.22)**	4.48 (0.81)	2.41 (0.64)	18.16 (1.90)*
Repeated a grade	-5.05 (-1.07)	-3.58 (-0.43)	8.46 (0.79)	-18.83 (-1.41)
Much absence in past school year	0.64 (0.14)	1.90 (0.27)	-15.32 (-1.41)	-2.78 (-0.25)
Expects post-high school education (base)	0.01 (0.00)	-4.87 (-0.61)	-7.55 (-0.77)	6.28 (0.57)
Reading pre-test	0.75 (20.30)**	0.66 (10.94)**	0.77 (11.47)**	0.65 (6.65)**
Enjoys reading (base)	-3.13 (-1.65)*	-8.09 (-2.07)**	-0.93 (-0.30)	1.81 (0.33)

Table D-5c (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT			
	Blacks	Hispanic	Asian	Whites & Others
Boston	8.55 (1.43)			16.97 (0.97)
Fresno	11.59 (1.85)*	13.79 (1.51)		20.28 (0.92)
Portland	7.07 (1.23)		34.59 (3.61)**	32.87 (1.91)*
San Diego	-3.07 (-0.48)	20.79 (1.93)*	13.14 (1.67)*	21.75 (0.78)
Sample size	674	175	174	121
F-test	35.96	14.13	16.62	7.04
R-squared	0.47	0.57	0.61	0.52
Adjusted R-squared	0.45	0.53	0.58	0.45

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-5d
Math Impacts by Race

Explanatory Variable	ESTIMATED COEFFICIENT			
	Blacks	Hispanic	Asian	Whites & Others
Constant term	29.08 (0.63)	-57.31 (-0.68)	123.03 (1.55)	228.36 (1.70)*
Treatment	11.03 (2.25)**	20.73 (2.16)**	6.61 (0.73)	14.64 (0.98)
Male	-19.43 (-3.92)**	-26.33 (-2.65)**	-0.67 (-0.07)	2.60 (0.17)
Receives AFDC and/or foodstamps	-3.50 (-0.66)	0.31 (0.03)	14.37 (1.34)	26.98 (1.71)*
Family size	0.33 (0.23)	-3.20 (-1.17)	-3.31 (-1.73)*	-2.55 (-0.61)
Lives without father	9.49 (1.80)*	-5.93 (-0.55)	-11.03 (-0.95)	-10.37 (-0.62)
Position in birth order	1.34 (1.35)	-5.51 (-2.63)**	0.69 (0.32)	-2.67 (-0.71)*
Problem with English		13.76 (0.84)	7.59 (0.82)	
Grade last year	10.71 (2.62)**	14.31 (1.91)*	10.83 (1.96)*	-0.05 (-0.00)
Repeated a grade	2.06 (0.34)	9.56 (0.85)	1.18 (0.08)	-16.54 (-0.92)
Much absence in past school year	-0.31 (-0.05)	5.48 (0.56)	17.03 (1.07)	30.97 (2.05)**
Expects post-high school education (base)	0.93 (0.16)	-2.27 (-0.21)	24.01 (1.70)*	7.67 (0.51)
Reading pre-test	0.28 (5.03)**	0.21 (2.15)**	0.31 (2.46)**	-0.22 (-1.37)
Math pre-test	0.53 (11.78)**	0.73 (8.10)**	0.44 (5.06)**	0.84 (6.65)**

Table D-5d (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT			
	Blacks	Hispanic	Asian	Whites & Others
Enjoys math (base)	-10.56 (-4.66)**	-11.58 (-2.46)**	-16.90 (-3.67)**	0.38 (0.06)
Boston	5.13 (0.65)			25.37 (1.09)
Fresno	1.69 (0.20)	22.45 (1.80)**		23.34 (0.78)
Portland	13.51 (1.80)*		-1.15 (0.08)	39.57 (1.70)*
San Diego	4.53 (0.53)	25.73 (1.72)*	14.85 (1.27)	69.69 (1.84)*
Sample size	674	176	174	121
F-test	34.08	14.17	12.52	5.71
R-squared	0.47	0.59	0.56	0.49
Adjusted R-squared	0.46	0.55	0.52	0.40

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-5e
Reading and Math Impact by Sex

Explanatory Variable	ESTIMATED COEFFICIENT			
	READING		MATH	
	Male	Female	Male	Female
Constant term	139.67 (3.37)**	113.27 (3.47)**	84.65 (1.53)	-6.02 (-0.14)
Treatment	0.45 (0.10)	11.37 (3.19)**	12.04 (1.98)**	12.41 (2.68)**
Black	-4.85 (-0.65)	-3.60 (-0.58)	-14.51 (-1.43)	-4.11 (-0.51)
Hispanic	-9.56 (-1.02)	-1.37 (-0.17)	-14.38 (-1.13)	-4.18 (-0.41)
Asian	20.20 (1.90)*	18.05 (2.01)**	30.52 (2.08)**	24.40 (2.05)**
Receives AFDC and/or foodstamps	0.35 (0.07)	-1.41 (-0.36)	3.33 (0.51)	-1.48 (-0.29)
Family size	-0.91 (-0.74)	-0.24 (-0.25)	-0.77 (-0.46)	-1.52 (-1.22)
Lives without father	-0.35 (-0.07)	9.83 (2.54)**	0.71 (0.10)	5.08 (1.01)
Position in birth order	-1.22 (-1.25)	-0.67 (-0.93)	-0.18 (-0.14)	0.66 (0.70)
Problem with English	-3.63 (-0.41)	-7.15 (-0.85)	8.34 (0.69)	2.41 (0.22)
Grade last year	7.72 (2.14)**	5.49 (2.05)**	9.57 (1.93)*	10.83 (3.12)**
Repeated a grade	-1.62 (-0.30)	-8.26 (-1.77)*	1.92 (0.26)	4.74 (0.78)
Much absence in past school year	-5.67 (-1.07)	2.25 (0.54)	4.73 (0.66)	3.83 (0.70)
Expects post-high school education (base)	1.44 (0.27)	-2.92 (-0.72)	1.73 (0.23)	3.43 (0.65)

Table D-5e (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT			
	READING		MATH	
	Male	Female	Male	Female
Reading pre-test	0.70 (17.00)**	0.76 (21.45)**	0.20 (2.88)**	0.28 (5.20)**
Enjoys reading (base)	-2.48 (-1.07)	-2.42 (-1.35)		
Math pre-test			0.57 (10.84)**	0.60 (13.10)**
Enjoys math (base)			-14.71 (-5.01)**	-6.99 (-3.35)**
Boston	15.08 (1.94)*	7.31 (1.25)	8.51 (0.80)	-1.46 (-0.19)
Fresno	24.52 (3.13)**	8.75 (1.40)	7.90 (0.73)	6.59 (0.80)
Portland	19.85 (2.74)**	12.35 (2.12)**	17.68 (1.78)*	7.31 (0.96)
San Diego	7.61 (1.04)	8.49 (1.46)	8.42 (0.82)	14.82 (1.87)*
Sample size	581	564	581	564
F-test	21.71	36.85	26.04	39.16
R-squared	0.42	0.56	0.48	0.59
Adjusted R-squared	0.40	0.55	0.46	0.58

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-6
General Knowledge of Contraception
Overall and By Sex

Explanatory Variable	ESTIMATED COEFFICIENT		
	All	Male	Female
Constant term	-1.80 (-1.73)*	-2.64 (-1.70)*	-0.60 (-0.42)
Treatment	2.55 (18.60)**	2.39 (11.15)**	2.61 (14.40)**
Male	0.12 (0.78)		
Age	-0.15 (-1.10)	0.05 (0.22)	-0.26 (-1.47)
Black	0.26 (1.09)	0.76 (2.02)**	-0.17 (-0.54)
Hispanic	0.48 (0.16)	0.41 (0.90)	-0.18 (-0.46)
Asian	-0.95 (-2.73)**	-0.51 (-0.92)	-1.30 (-2.84)**
Receives AFDC and/or foodstamps	-0.35 (-2.32)**	-0.38 (-1.65)*	-0.26 (-1.33)
Family size	0.03 (0.88)	0.05 (0.84)	0.01 (0.24)
Lives without father	0.18 (1.20)	0.21 (0.89)	0.16 (0.80)
Position in birth order	-0.03 (-1.03)	-0.04 (-0.97)	-0.02 (-0.50)
Problem with English	0.55 (1.84)*	0.52 (1.21)	0.75 (1.76)*
Much absence in past school year	-0.17 (-1.04)	-0.10 (-0.41)	-0.29 (-1.35)
Expects post-high school education (base)	-0.14 (-0.82)	-0.27 (-1.01)	0.01 (0.06)

Table D-6 (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT		
	All	Male	Female
Reading pre-test	0.01 (4.28)**	0.01 (3.60)**	0.00 (2.11)**
Ever had sex	0.12 (0.70)	0.13 (0.44)	0.12 (0.56)
Previously had sex education course	0.05 (0.35)	0.26 (1.15)	-0.15 (-0.76)
Woman responsible for birth control (base)	-0.46 (-2.76)**	-0.30 (-1.10)	-0.60 (-2.89)**
Birth control avail- ability score (base)	0.16 (3.76)**	0.13 (1.99)**	0.20 (3.33)**
Knowledge of birth control score (base)	0.32 (9.02)**	0.23 (4.07)**	0.41 (8.62)**
Best age to have children (base)	0.02 (0.81)	0.00 (0.08)	0.03 (1.15)
Know consequences of pregnancy (base)	0.14 (1.38)	0.36 (2.13)**	-0.06 (-0.48)
Boston	0.18 (0.76)	-0.01 (-0.04)	0.32 (1.04)
Fresno	0.97 (3.94)**	0.93 (2.37)**	1.03 (3.25)**
Portland	0.28 (1.24)	0.60 (1.72)*	-0.05 (-0.17)
San Diego	0.73 (3.25)**	0.64 (1.83)*	0.75 (2.50)**
Sample size	937	451	486
F-test	29.90	11.29	22.41
R-squared	0.45	0.39	0.54
Adjusted R-squared	0.44	0.35	0.51

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-7

Knowledge of Contraceptive Availability
Overall and by Sex

Explanatory Variable	ESTIMATED COEFFICIENT		
	All	Male	Female
Constant term	-1.01 (-1.26)	-2.04 (-1.77)	-0.16 (-0.14)
Treatment	1.73 (16.42)**	1.67 (10.48)**	1.75 (12.15)**
Male	-0.17 (-1.48)		
Age	0.13 (1.21)	0.10 (0.60)	0.15 (1.07)
Black	-0.33 (-1.77)*	-0.21 (-0.74)	-0.37 (-1.50)
Hispanic	0.02 (0.09)	0.22 (0.65)	-0.02 (-0.07)
Asian	-0.81 (-3.05)**	-0.82 (-2.03)**	-0.78 (-2.14)
Receives AFDC and/or foodstamps	-0.24 (-2.07)**	0.00 (-0.03)	-0.41 (-2.60)**
Family size	-0.00 (-0.00)	0.01 (0.29)	-0.03 (-0.72)
Lives without father	0.03 (0.23)	-0.16 (-0.88)	0.19 (1.24)
Position in birth order	0.00 (0.08)	0.06 (1.89)*	-0.05 (-1.70)*
Problem with English	0.01 (0.06)	-0.05 (-0.15)	0.20 (0.58)
Absent much during past school year	-0.09 (-0.71)	-0.06 (-0.34)	-0.15 (-0.90)

Table D-7 (cont'd)

Explanatory Variable	ESTIMATED COEFFICIENT		
	All	Male	Female
Expects post-high school education (base)	0.02 (0.16)	0.19 (0.94)	-0.14 (-0.83)
Reading pre-test	0.00 (3.93)**	0.00 (3.39)**	0.00 (2.25)**
Ever had sex	0.03 (0.25)	0.04 (0.17)	0.01 (0.04)
Previously had sex education course	0.20 (1.70)*	0.27 (1.62)	0.05 (0.30)
Woman responsible for birth control (base)	-0.28 (-2.19)**	-0.19 (-0.93)	-0.37 (-2.24)**
Birth control availability score (base)	0.29 (8.65)**	0.28 (5.56)**	0.31 (6.56)**
Knowledge of birth control score (base)	0.13 (4.57)**		0.16 (4.14)**
Best age to have children (base)	0.01 (0.43)	-0.01 (-0.52)	0.02 (1.00)
Know consequences of pregnancy (base)	0.06 (0.78)	0.20 (1.58)	-0.07 (-0.66)
Boston	0.04 (0.22)	-0.13 (-0.46)	0.11 (0.50)
Fresno	0.36 (1.92)*	0.58 (1.99)**	0.12 (0.49)
Portland	0.44 (0.25)	0.36 (1.38)	-0.29 (-1.22)
San Diego	0.55 (3.18)**	0.68 (2.63)**	0.41 (1.74)*
Sample size	937	451	486
F-test	27.04	11.29	17.64
R-squared	0.43	0.39	0.48
Adjusted R-squared	0.41	0.35	0.45

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.

Table D-8

Pooled Probit Estimation of Participation in
Sexual Intercourse During the Previous Two Months

<u>Explanatory Variable</u>	<u>Coefficient</u>
Constant term	-1.44 (-2.75)**
Treatment	-0.05 (-0.42)
Male	0.68 (6.00)**
Age	-0.05 (-0.48)
Black	0.47 (2.45)**
Hispanic	0.32 (1.38)
Asian	-0.41 (-1.23)
Receives AFDC and/or foodstamps	0.08 (0.68)
Family size	-0.01 (-0.19)
Lives without father	-0.08 (-0.63)
Position in birth order	-0.05 (-1.96)**
Problem with English	-0.34 (-0.91)
Absent much during past school year	0.08 (0.59)
Expects post-high school education (base)	-0.01 (-0.06)
Previously had sex education course	0.19 (1.50)

Table D-8 (cont'd)

Explanatory Variable	Coefficient
Woman responsible for birth control (base)	0.11 (0.81)
Birth control availability score (base)	0.00 (0.12)
Knowledge of birth control score (base)	0.04 (.28)
Best age to have children (base)	-0.02 (-1.51)
Know consequences of pregnancy (base)	0.06 (0.75)
Boston	0.34 (1.83)*
Fresno	0.04 (0.20)
Portland	0.25 (1.42)
San Diego	-0.09 (-0.46)
Sex during previous two months (base)	1.60 (12.71)**

Value of likelihood function for restricted model = 553.47

Value of likelihood function for unrestricted model = 337.28

Chi-squared = 432.18** with 24 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

Table D-8a

Probit Estimation for Sexual Intercourse Participation
of Females During the Previous Two Months

Explanatory Variable	Coefficient
Constant term	-1.88 (-2.17)**
Treatment	0.25 (1.41)
Age	0.13 (0.75)
Black	0.44 (1.42)
Hispanic	0.56 (1.49)
Asian	0.03 (0.06)
Receives AFDC and/or foodstamps	0.14 (0.71)
Family size	0.03 (0.68)
Lives without father	0.17 (0.89)
Position in birth order	-0.09 (-2.19)
Problem with English	-2.60 (-0.20)
Absent much during past school year	0.26 (1.30)
Expects post-high school education (base)	-0.02 (-0.12)
Previously had sex education course	-0.14 (-0.67)
Woman responsible for birth control (base)	-0.04 (-0.19)

Table D-8a (cont'd)

Explanatory Variable	Coefficient
Birth control availability score (base)	0.05 (0.98)
Knowledge of birth control score (base)	0.07 (1.56)
Best age to have children (base)	-0.04 (-1.61)
Know consequences of pregnancy (base)	0.18 (1.36)
Boston	0.40 (1.42)
Fresno	-0.31 (-0.99)
Portland	0.24 (0.84)
San Diego	0.23 (0.76)
Sex during previous two months (base)	1.73 (8.10)**

Value of likelihood function for restricted model = 214.24
 Value of likelihood function for unrestricted model = 136.81
 Chi-squared = 154.86 with 23 degrees of freedom

** Significant at the .05 level.

Table D-8b

Probit Estimation for Sexual Intercourse Participation
of Males During the Previous Two Months

Explanatory Variable	Coefficient
Constant term	-0.58 (-0.84)
Treatment	-0.26 (-1.70)**
Age	-0.16 (-1.04)
Black	0.64 (2.48)**
Hispanic	0.24 (0.75)
Asian	-0.71 (-1.52)
Receives AFDC and/or foodstamps	0.03 (0.20)
Family size	-0.02 (-0.45)
Lives without father	-0.21 (-1.21)
Position in birth order	-0.03 (-0.91)
Problem with English	-0.06 (-0.14)
Absent much during past school year	-0.12 (-0.68)
Expects post-high school education (base)	-0.03 (-0.15)
Previously had sex education course (base)	0.31 (1.87)*
Woman responsible for birth control (base)	0.17 (0.89)

Table D-8b (cont'd)

Explanatory Variable	Coefficient
Birth control available score (base)	-0.04 (-0.81)
Knowledge of birth control score (base)	-0.00 (-0.03)
Best age to have children (base)	-0.01 (-0.46)
Know consequences of pregnancy (base)	0.02 (0.13)
Boston	0.30 (1.15)
Fresno	0.30 (1.09)
Portland	0.34 (1.37)
San Diego	-0.19 (-0.73)
Sex during previous two months (base)	1.53 (9.23)**

Value of likelihood function for restricted model = 296.51
 Value of likelihood function for unrestricted model = 183.49
 Chi-squared = 266.04** with 23 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

Table D-9

Probit Estimation of the Use of Birth Control
During Participants' Last Sexual Intercourse

<u>Explanatory Variable</u>	<u>Coefficient</u>
Constant term	-1.99 (-1.55)
Treatment	0.20 (0.77)
Male	-0.10 (-0.30)
Age	0.15 (0.56)
Black	0.40 (0.99)
Hispanic	0.69 (0.33)
Asian	0.23 (0.23)
Receives AFDC and/or foodstamps	-0.37 (-1.30)
Family size	-0.20 (-2.08)**
Lives without father	-0.28 (-0.94)
Position in birth order	-0.02 (-0.32)
Problem with English	0.28 (0.23)
Absent much during past school year	-0.07 (-0.24)
Expects post-high school education (base)	-0.05 (-0.19)
Previously had sex education course	0.41 (1.40)
Woman responsible for birth control (base)	0.74 (2.41)**

Table D-9 (cont'd)

Explanatory Variable	Coefficient
Birth control availability score (base)	0.05 (0.60)
Knowledge of birth control score (base)	0.06 (0.95)
Best age to have children (base)	0.04 (1.24)
Know consequences of pregnancy (base)	-0.01 (-0.04)
Contraceptive use meaning 1 (base)	1.86 (6.54)**
Boston	0.29 (0.75)
Fresno	0.26 (0.53)
Portland	-0.13 (-0.35)
San Diego	-0.29 (-0.61)

Value of likelihood function for restricted model = 114.64

Value of likelihood function for unrestricted model = 75.66

Chi-squared = 77.95** with 24 degrees of freedom

** Significant at the .05 level.

Table D-10

Probit Estimation for Learning Sample Membership

Explanatory Variable	Coefficient
Constant term	-0.51 (4.53)**
Male	-0.20 (-2.89)**
Hispanic	0.10 (0.80)
Black	0.17 (1.59)
Asian	0.35 (2.62)**
Age	0.01 (0.90)
Treatment	0.02 (0.28)

Value of Likelihood Function for Restricted Model = 947.36
 Value of Likelihood Function for Unrestricted Model = 939.36
 Chi-squared = 16.00* with 6 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

Table D-11

Probit Estimation for Life Skills Sample Membership

Explanatory Variable	Coefficient
Constant term	0.32 (2.96)**
Male	-0.03 (-4.69)*
Hispanic	0.00 (0.01)
Black	0.10 (0.94)
Asian	0.23 (1.80)*
Age	-0.03 (-0.43)
Treatment	-0.03 (-0.49)

Value of Likelihood Function for Restricted Model = 1080.30
 Value of Likelihood Function for Unrestricted Model = 1066.30
 Chi-squared = 27.40** with 6 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

Table 12

Probit Estimation for Dropping Out of the Program

Explanatory Variable	Coefficient
Constant term	-0.73 (-5.98)**
Male	0.09 (1.24)
Hispanic	-0.19 (-1.43)
Black	-0.29 (-2.55)**
Asian	-0.44 (-3.00)**
Age	-0.04 (-0.54)
Treatment	0.05 (0.73)

Value of likelihood function for restricted model = -744.84

Value of likelihood function for unrestricted model = -738.57

Chi-squared = 12.54* with 6 degrees of freedom

* Significant at the .10 level.

** Significant at the .05 level.

Table D-13

Multinomial Logit for Knowledge of
the Consequences of Pregnancy

Explanatory Variable	Y = 01	Y = 02
Constant Term	3.01 (0.15)	-0.70 (-0.03)
Male	0.20 (0.38)	-0.07 (-0.07)
Age	0.22 (0.45)	0.08 (0.17)
Black	-6.36 (-0.31)	-6.31 (-0.31)
Hispanic	-6.86 (-0.34)	-6.67 (-0.33)
Asian	-6.14 (-0.30)	-6.39 (-0.31)
Receives AFDC and/or foodstamps	0.26 (0.49)	0.32 (0.60)
Family size	0.04 (0.27)	0.04 (0.33)
Lives without father	-0.61 (-1.17)	-0.66 (-1.24)
Position in birth order	-0.01 (-0.12)	0.04 (0.41)
Problem with English	-0.01 (-0.01)	0.20 (0.20)
Absent much in past school year	-0.29 (-0.54)	-0.30 (-0.54)
Expects post-high school education (base)	-0.40 (-0.70)	-0.20 (-0.33)
Ever had sex	-0.76 (-1.30)	-0.53 (-0.88)
Previously had sex education course	0.15 (0.30)	0.31 (0.60)

Table D-13 (cont'd)

Explanatory Variable	Y = 01	Y = 02
Reading pre-test	0.00 (0.77)	0.01 (1.09)
Woman responsible for birth control (base)	-0.31 (-0.60)	-0.47 (-0.89)
Birth control availability score (base)	-0.12 (-0.80)	-0.10 (-0.62)
Knowledge of birth control score (base)	0.17 (1.33)	0.18 (1.40)
Best age to have children (base)	0.12 (1.65)*	0.15 (2.14)**
Know consequences of pregnancy (base)	0.89 (2.33)**	1.89 (4.78)**
Boston	1.57 (1.59)	1.00 (1.00)
Fresno	0.26 (0.33)	-0.11 (-0.13)
Portland	0.64 (0.75)	0.58 (0.67)
San Diego	0.01 (0.13)	-0.25 (-0.33)

Value of Likelihood Function for Restricted Model = -531.06
 Value of Likelihood Function for Unrestricted Model = -529.81
 Chi-Squared = 2.5 with 2 degrees of freedom.

* Significant at the .10 level.

** Significant at the .05 level.

NOTE: Numbers in parentheses are T-statistics.